TAMPERE UNIVERSITY OF TECHNOLOGY

ENRIQUE SÁNCHEZ-MONGE FERNÁNDEZ

IMAGE RECONSTRUCTION WITH MISSING WEDGE CORRECTION IN ELECTRON TOMOGRAPHY

Master of Science Thesis

Examiner: Professor Ulla Ruotsalainen
Examiner and topic approved in
the Computing and Electrical
Engineering Faculty Council
meeting on 12 January 2011

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master of Science Degree Programme in Information Technology
SÁNCHEZ-MONGE FERNÁNDEZ, ENRIQUE: Image Reconstruction with Missing Wedge Correction in Electron Tomography
Master of Science Thesis, 85 pages, 5 Appendix pages
June 2011
Major: Signal processing
Examiner: Professor Ulla Ruotsalainen
Instructors: Assistant Prof. Sari Peltonen, Uygar Tuna, MSc.
Keywords: TEM, electron tomography, image reconstruction, missing wedge, sinogram, compressed sensing

Electron Tomography (ET) is a technique that makes possible to obtain 3D reconstructions of organic and inorganic materials from Transmission Electron Microscope (TEM) images. However, physical limitations in the acquisition system leave a blind region, better known as the missing wedge, which causes anisotropic resolution and artifacts in the reconstructed tomograms. In this work, the effects of the missing wedge were studied and different methods were proposed in order to correct them. In addition, we propose an iterative reconstruction method with missing wedge correction for ET, improving results compared with traditional reconstructions and reducing the effects due to the missing wedge. The use of an adaptive filter for regularization during reconstruction does not rely on *a priori* assumptions, making it appropriate for reconstructing TEM tomograms, where the complexity of biological specimens makes it difficult to use strong constraints and *a priori* information to improve reconstructions. The proposed method outperformed traditional reconstructions such as Filtered Back-Projection and was able to obtain a reconstruction where the artifacts caused by the missing wedge were not detected. The high presence of noise in real TEM data leads to future work in denoising of projections to improve the reconstruction performance.

PREFACE

This thesis work was part of the BIQ Project (Biomedical Image Quantification), and was funded by the University Alliance Finland, composed of Tampere University of Technology, University of Jyväskylä and University of Tampere. It was accomplished in the Methods and Models for Biological Signals and Images (M^2oBSI) research group, at the Department of Signal Processing, Tampere University of Technology, Finland, between June 2010 and June 2011. The work was performed under the supervision of Professor Ulla Ruotsalainen, Assistant Professor Sari Peltonen and M.Sc. Uygar Tuna.

I am really thankful to my supervisors Ulla, Sari and Uygar, for advice and support. I would like to thank to Professor Holland Cheng from the department of Molecular & Cellular Biology, at the University of California Davis for the advice received. Also, I wish to show my gratitude to Dr. Varpu Marjomäki and M.Sc. Lassi Paavolainen from the department of Biological and Environmental Science at University of Jyväskylä for the collaboration and help.

I would like to give special thanks to Alessandro Foi for his time and guiding, and I would also like to thank my colleagues in the M^2oBSI group and especially to Uygar Tuna for all the time, help and review given.

Finally, I wish to thank Ana María, our foster cat Lyly and my family for company, support and love during the developing of this work.

Enrique Sánchez-Monge Fernández
Tampere, June 2011

Finninmäenkatu 4 H 84
33710 Tampere, FINLAND
+358 466 117 678
enrique.sanchez-mongefernandez@tut.fi

TABLE OF CONTENTS

ABBREVIATIONS, NOTATION AND SYMBOLS

$\delta(x)$	Discrete Impulse Funcion
Δt	Time Increment
e_{yz}	Elongation Factor in the yz-plane
$f(x,y)$, $I(i,j)$	Two-dimensional Image
$F(u,v)$	2D Fourier Transform of $f(x,y)$
$F(\omega,\theta)$	2D Fourier Transform of $f(x,y)$, expressed in Polar Coordinates
Φ	Filtering Operator
$g_{\theta_0}(\rho)$	Projection Vector at Angle θ_0
$g(\rho,\theta)$	Sinogram
$G(\omega,\theta)$	One-Dimensional Fourier transform of $g(\rho,\theta)$ with respect to ρ
I_{xx}	Second Order Derivative of I with respect to x
$\overrightarrow{\delta L}$	Inpainting Measure of Change of Information
∇	Gradient Operator
$\overrightarrow{N}$	Inpaiting Propagation direction
$O(\cdot)$	Computational Complexity
Ω	Inpainting Gap
$\perp$	Perpedincular
$P(k,\lambda)$	Poisson Distribution
$\mathbb{R}^m$	m-Dimensional Set of Real Numbers
$\cup$	Set Union
T	Transform Operator
T^{-1}	Inverse Transform Operator
y	Image Spectrum defined by Transform Operator T

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
ART	Algebraic Reconstruction Technique
BF	Bright Field
BM3D	Block Matching 3D Filter
CCD	Charge-Coupled Device
CET	Cryo-Electron Tomography
CFS	Constrained Fourier Space
CS	Compressed Sensing, Compressive Sampling
CT	Computed Tomography
CTF	Contrast Transfer Function
DCT	Discrete Cosine Transform
DFR	Direct Fourier Reconstruction
DFT	Discrete Fourier Transform
EM	Electron Microscopy
ET	Electron Tomography
FBP	Filtered Back-Projection
FT	Fourier Transform
HAADF	High Angle Annular Dark Field
MLEM	Maximum Likelihood Expectation Maximization
MRP	Median Root Prior
MSE	Mean Squared Error
OSEM	Ordered Subsets Expectation Maximization
PET	Positron Emission Tomography
POCS	Projections Onto Convex Sets
SART	Simultaneous Algebraic Reconstruction Technique
SIRT	Simultaneous Iterative Reconstructions Technique
SPECT	Single-Photon Emission Computed Tomography
SNR	Signal-to-Noise Ratio
STEM	Scanning Transmission Electron Microscope
TEM	Transmission Electron Microscope
WBP	Weighted Back-Projection

1. INTRODUCTION

In his work *Electron Tomography* , Joachim Frank gives a good definition for electron tomography as "any technique that employs the transmission electron microscope to collect projections of an object that is tilted in multiple directions and uses these projections to reconstruct the object in its entirety." [1]. That is, three-dimensional (3D) structures are obtained by using two-dimensional (2D) data, which are projections obtained at different points of view with the Transmission Electron Microscope (TEM). This technique is not new at all, as it has been used for approximately 40 years, since De Rosier and Klug established the principles of 3D reconstruction from electron microscope images, which are also known as micrographs [2]. Although they outlined the conditions for any 3D structure, they exploited the symmetry property of the particles to reconstruct the original structure with few micrographs, and in some cases even just one projection was enough. This is not the case, however, for asymmetric biological specimens where neither redundancies nor symmetries can be exploited. Nevertheless, the methods for obtaining tomographic reconstruction based on micrographs were established and have been developed until today.

At the moment, high-resolution electron microscopes can generate images close to atomic resolution, meaning an important improvement for biological and material sciences [3]. By using electron tomography, it is possible to obtain much better resolutions than by using optical microscopy. Thus, it is normal to attain a 3D structural description of a specimen at a usual resolution which is in the order of $5-10$ nm [4, 5]. Moreover, nowadays it is not unusual to speak of resolutions better than 1 nm, down to 0.2 nm [6], not to mention that the field of view can be of hundreds of nm [3]. However, when dealing with biological samples, their sensitivity to radiation becomes a serious problem, where over-exposure to the electron beam incinerates the sample [7]. In order to optimize resolution, as many images as possible from the widest angular range should be taken. However, this contradicts the requirement of maintaining low electron doses to avoid damages [8, see 4].

The description of the idea behind electron tomography is simple. An object is imaged over a full range of tilts using the TEM, and these images are then used to obtain a reconstruction of the original object. Nevertheless, this approach presents a problem in electron microscopy: it is not possible to obtain a full tilt range of projections. Physical limitations of the holder that contains the sample and the

increment of traversed path of electrons at high tilting angles restrict the maximum possible tilt angle. This reduces the range of projections thus creating a wedge shaped region in Fourier space known as the *missing wedge*. In this manner, covered angular range in commercial microscopes does not commonly go beyond $\pm 60^\circ$ [1]. Hence the motivation for this work is to compensate, or fill this missing wedge and perform the reconstruction of the compensated projection data as good as possible to an image of the original object (the specimen).

This work is divided in seven chapters. After the introduction, Chapter 2 introduces the concepts of electron tomography and in transmission electron microscopy, and then establishes the problem of the missing wedge. In Chapter 3, image reconstruction from projections is explained and the most frequently used methods are described. The methods to overcome the missing wedge problem are then described in Chapter 4. Chapter 5 presents the data sets used in the simulations and reconstructions, after which in Chapter 6 the obtained results are shown. Finally, the results are discussed and final conclusions are drawn in Chapter 7.

2. ELECTRON TOMOGRAPHY

Electron tomography (ET) has become an important field in biological and material sciences since it was firstly conceived half a century ago. It constitutes an essential tool for visualizing 3D internal structures of biological, and non-biological, specimens down to nanometric resolution. In this chapter ET and its underlying principles are described in order to make the study performed for this thesis understandable. It starts with the basics of the imaging technique used in ET, transmission electron microscopy, followed by the principles of ET.

2.1 Transmission Electron Microscopy (TEM)

Electron tomography relies on tomographic reconstruction using micrographs, projection images obtained from the Transmission Electron Microscope (TEM). Therefore, a brief introduction to transmission electron microscopy is needed in order to understand completely the subjects dealt in this work.

A first step to understand the TEM is to know what microscopy is and why it is needed. Microscopy can be defined as the field of science that deals with the magnification of objects or samples that cannot be seen by naked human eye. The smallest object, which unaided human eye can distinguish, is given by the spatial resolution of the eye, that is approximately 0.1 mm [9, p. 5]. Therefore, in order to see beyond this resolution, microscopes are needed. To achieve this, microscopes use properties of light, in optical microscopy, and electrons in electron microscopy. Electron microscopy poses a very useful approach since the electrons have a wavelength comparable to atomic dimensions, 0.17 nm at 50 V and 5 pm at 50 kV [6]. This fact allows us to image much smaller objects than the samples in optical microscopy.

A special type of electron microscopes is the TEM. This family of microscopes uses high energy electron beams that penetrate a thin specimen interacting with its atoms, and due to these interactions some electrons are deflected, or scattered, changing their trajectories. These interactions between electrons and matter will be determined by the structure of the latter. In this manner, scattered electrons carry information of the specimen which is then used to create contrast and thus obtain an image of the sample. The transmitted electron beam is then focused using magnetic lenses on an image plane and the image is recorded. The properties and

high magnification offered by TEM make them widely used, especially, in the fields of life science and material science.

2.1.1 Acquisition System

Extensive literature regarding to the physical principles under electron microscopy exists, for example see [10–12], but a detailed description of the physical principles falls out of the scope of this work. However, it is necessary to give a brief introduction to transmission electron microscopy in order to make the reader familiar with the principles behind the system. Roughly, the TEM can be divided into three sections (or stages) [13]:

1. Illumination system.
2. Specimen stage.
3. Imaging system.

1. Illumination system.

This is the stage where the beam of electrons is generated. There are basically three different types of electron guns, based on three different principles, *thermionic emission*, *Schottky emission* and *field emission*. Since the images used in this work are taken from a field emission TEM, see Figure 2.1, only this type of microscope is described. In a field emission gun, a thin beam of electrons, typically between 0.1 and 5 nm in diameter [9], is generated at a cathode, which consists of a sharp tip, where a high voltage is applied generating a strong electric field enough to allow electrons to jump out of the tip.

After emission, electrons are accelerated up to the desired energy by applying a strong potential difference between the cathode and the anode (200 kV for the images used in this work), and focused onto the specimen by a set of lenses. This is known as the *accelerating voltage* of the microscope, and the higher it is, the higher energy the electrons acquire. Higher energies decrease the wavelength of the electrons [6, 9], allowing the improvement in the resolution, and also the increase in the specimen thickness. However, if high accelerating voltage is used, although it will increase the *mean free path* of the

Figure 2.1: *The Field Emission TEM used to take the images in this work, model JEM-2100F [14].*

electron[1] , it would also destroy the biological sample before any image could be taken [6]. This is because the electron beam is a form of high energy ionizing radiation and chemically modifies and even destroys biological samples.

2. Specimen stage.

Interaction of electrons with matter occurs at this stage. There are several types of interactions, yielding different signals that are used in different kinds of microscopy [6, 9, 15]. The most important interaction in TEM is scattering. Electrons from the incident beam hit the specimen and are deviated from their original trajectories as they interact with the atoms. This scattering process creates contrast in the recorded image, Figure 2.2. There are different types of scattering, and they are not explained in detail in this thesis as they fall outside the field of this work. However it is necessary to differentiate between *single scattering* and *multiple scattering*. The former is present when an electron suffers a single scattering event, or if it does not suffer any scattering at all [15]. Single scattering is the desirable process as it is easier to interpret the image obtained in the TEM. If the electrons suffer more than approximately twenty scattering events, it is denominated by multiple scattering and it can be assumed that it is not present in TEM images unless the specimen is considerably thicker than the mean free path [15]. Moreover, scattering events can be divided into two more types, known as *elastic* and *inelastic scattering*. Electrons are scattered *inelastically* if there is an energy transfer between the electron and the atom that causes the scattering [15,16]. On the other hand, if there is not any appreciable energy transfer, it is said to be elastic scattering.

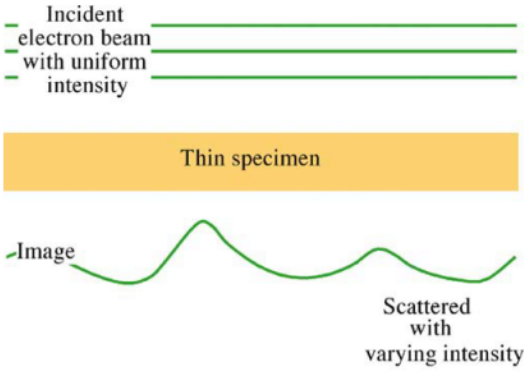


Figure 2.2: *The electrons that traverse the specimen are scattered. This changes the spatial distribution of the incident beam, which is observed as intensity changes in the image [15].*

The probability that an electron is scattered is measured by the *scattering cross-section*. The scattering cross-section depends on the incident beam energy and the atomic number of the atom that causes the scattering. It also varies with the scattering angle, giving information of how likely an electron will be scattered with a higher angle than a given angle α [16]. The acceptance angle of the aperture in the microscope determines which scattered electrons reach the imaging system.

[1] The mean free path is the average distance that an electron travels through the sample between scattering events [15].

This discriminates high scattering angles from the low ones and generates contrast from these differences, assuming that single scattering is mostly present [16]. This is known as *amplitude contrast*, and if multiple scattering is present, the interpretation of the image becomes more difficult and then the reconstruction result is degraded. A second type of contrast is the *phase contrast*, where the contrast comes from the phase shifts between unscattered electrons and the electrons which suffered elastic scattering [17].

3. Imaging system.

Once electrons traverse the specimen, they are focused by a strong lens into the image plane, consisting of a *scintillator* or *phosphor screen*, which emit visible light when hit by electrons. In the scintillator electrons are converted into photons and then are redirected towards a charge-coupled device (CCD) where the image is recorded. This is usually performed by using fiber optic or lenses [17]. These optical lenses allow the photons to be further focused onto the CCD [18]. The advantages of CCD sensors are the large dynamic range and the high linearity of the signal with respect to the incoming electrons. However, the CCD sensors exhibit poor performance in high frequency and with higher accelerating voltages [18]. Another disadvantage arises from the relatively small sensor size, which leads to a compromise between the desired magnification and its field of view [17]. In this way, if high resolution is desired, only a small part of the specimen should be imaged. On the other hand, if a greater area is desired to be imaged, poor resolution is obtained. However, the use of CCD sensors allows the use of automated image acquisition, thus reducing the electron dose and improving the acquisition stage.

2.1.2 Radiation Damage

An important problem in ET is the damage by excessive electron doses in the sample that will degrade the specimen, and in some cases it may even destroy the sample. This is known as *radiation damage* or *beam damage*. In particular, since biological samples are sensitive to radiation, over-exposure to the electron beam should be avoided. Moreover, if an excessive electron dose is applied on the sample, quantitative 3D reconstructions become questionable [1].

At room temperature, for the plastic embedded samples, the main effect of radiation damage is mass loss [18]. The sample shrinks mostly in the direction of the beam. This poses a great problem in 3D reconstruction, where successive images are recorded at different tilt angles. If at each recorded image the sample is shrunk, then the different images do not constitute projections of exactly same object, and the reconstruction is compromised. Beam damage is reduced in cryo-Electron To-

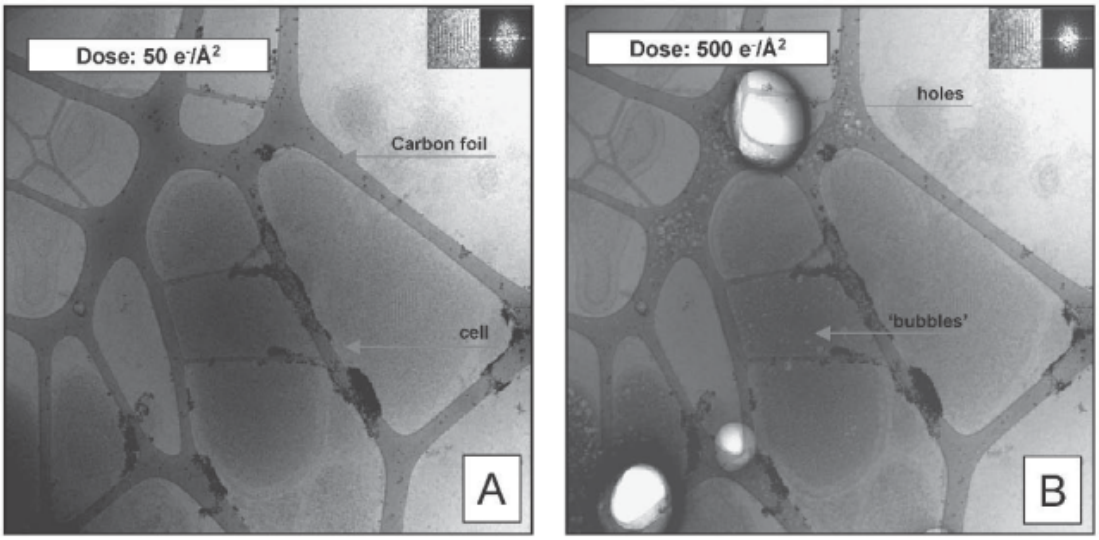


Figure 2.3: *Effects of beam damage over cryo-electron micrographs of an ice embedded prokaryotic cell [7]. Higher electron doses will degrade the specimen (right) where bubbles and holes appear.*

mography (CET), where the sample has been frozen at low temperatures, and only gradual loss of high resolution and occasionally bubbles are present [18], see Figure 2.3. Nevertheless, in CET, electron dose must be kept as low as possible, and usually the angular increment in the tilt series is larger, This angular increment is usually two degrees or more, in comparison with plastic embedded samples, where increment is normally one or two degrees [19].

2.1.3 Noise in the Transmission Electron Microscope

Selecting the appropriate electron dose when imaging a sample is a substantial problem faced in ET. Low electron doses will yield noisy images while higher electron doses will result in less noisy images but will degrade the specimens, and even destroy them. Thus, noise characterization becomes an important task.

In his PhD thesis and further works [20, 21], Frangakis mentions two principal noise sources, the illuminating electron beam and the image recording system, the former being the biggest contribution to the overall noise. The noise generated by the electron beam is concretely quantum noise, originated by the discrete character of electrons and the random fluctuations in the recorded signal. Each recorded pixel will be hit by a random number of pixels following a Poisson distribution [20, 21],

$$P(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{2.1}$$

where k represents the number of electrons hitting a given pixel and λ represents the corresponding pixel of a "hypothetical image that has been recorded without

noise" [20]. Thus according to [20], the expected number of electrons hitting a given pixel i would be,

$$E\{k_i\} = \lambda_i, \tag{2.2}$$

and the covariance between pixels i and j,

$$\mathrm{Cov}\{k_i, k_j\} = \delta_{ij}\lambda_i \qquad \text{with } \delta_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \tag{2.3}$$

Consequently, following the covariance term, each pixel is statistically independent with respect to all other pixels with a variance depending on the signal level. Further assumptions can be done, provided that the signal of interest comes from small deviations from the mean [20], and the noise is determined by the intensity level. In this case additive independent noise can be assumed . Moreover, in addition with other noise sources, such as the CCD and multiple scattering, the overall noise in the recorded images can be assumed to be additive Gaussian noise [21].

2.2 Tomography Using Transmission Electron Microscopy

Once the fundamentals of TEM imaging have been discussed, the main framework of this thesis can now be presented. Electron tomography, or tomographic reconstruction using the TEM is described in the following subsections.

2.2.1 Principles of Electron Tomography

The goal of ET is to obtain a 3D reconstruction, known as tomogram, of the inner structure of a given specimen, widely used in biological samples, such as cells and intra cellular structures. It is performed by imaging the specimen over a series of different views using the TEM and then using these images to estimate the structure of the original specimen. This is possible because the TEM images of the specimen basically consist in 2D projections of the specimen's structure or density. In Figure 2.4 a simplified version of the process is illustrated.

The most simple acquisition system is known as single-axis tilting, where the specimen to be imaged is placed in the tip of a holder, see Figure 2.5, and it is tilted over the holder's axis using a goniometer. In this manner, different views of the specimen at different angles are obtained. Initially this procedure was performed manually, but this was a slow process and caused high exposure of the specimen to the electron beam [5, 18], leading to radiation damage. For this reason automated ET was developed, allowing image acquisition with low electron dose conditions, and correcting automatically possible changes in focus and shifts in the specimen [17,18]. The image acquisition is then performed over a wide range of angles, starting

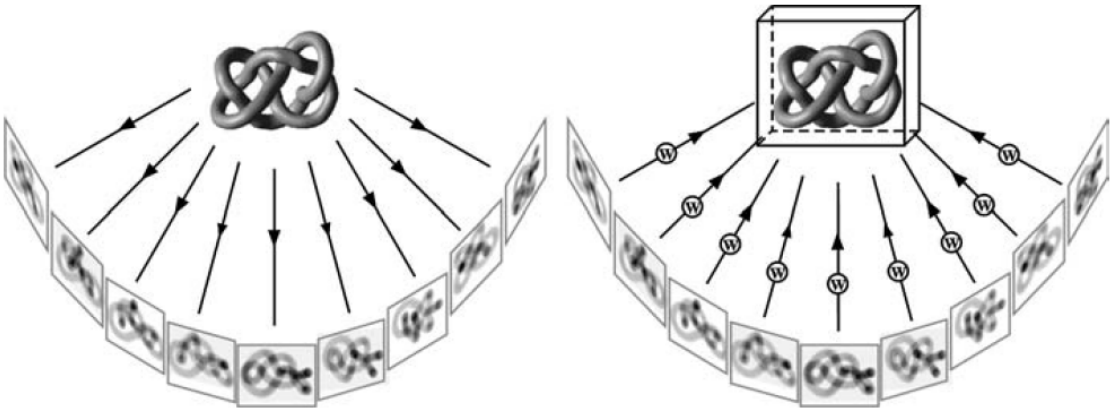


Figure 2.4: *Tomographic reconstruction principle [18]. Left: a set of different projection images of a specimen is acquired. Right: using these projection images the original specimen is reconstructed.*

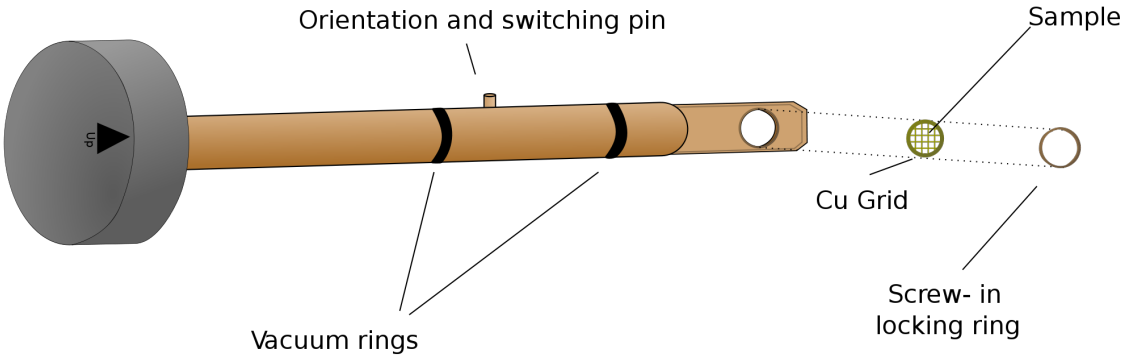


Figure 2.5: *Specimen holder [22].*

from zero tilt and progressively increasing the tilt angle by increments of 1° or 2°. Once the data set is finally obtained it can be used to perform a reconstruction of the tomogram using any of the available reconstruction methods. However, some conditions should be met if a faithful reconstruction of the original object is desired.

2.2.2 The Projection Requirement

Tomographic reconstruction is based on reconstruction principles that rely on projections. Thus, in order to be capable of reconstructing an object based on TEM images, it is essential that these represent projections through the structure of the sample. This is known as the *projection requirement* [3, 23]. In the case of TEM projection images, it is not as straightforward as in X-ray Computed Tomography (CT), where the projections are ray-sums of the absorption values along the path of the ray across the specimen. Interactions of electrons with the specimen in TEM are completely different than in X-ray CT and are not exactly ray-sums. However, the ray-sum condition of the projections is not the only possibility, as it is enough to be

a monotonically varying function of some physical property of the sample [16]. This is the case for some of the imaging techniques in TEM. As mentioned previously, there are two different types of imaging techniques, where contrast in the image comes from different phenomena, amplitude contrast and phase contrast.

The most used imaging technique is achieved by scattering contrast [16], also known as amplitude contrast. In this case the contrast comes from the *mass-thickness*[2] of the specimen, and this is especially predominant in stained samples. Higher mass-thickness will scatter the incoming electrons to larger angles, and if the scattering angle is large enough, the electron will be blocked by the aperture of the microscope, resulting in a dark contrast in the image. Regions with lower mass-thickness will not scatter the electrons that strongly, generating bright parts [3]. In particular, Hawkes establishes that the number of electrons per unit area hitting a pixel is proportional to $\exp(-\mu/\mu_i)$, where μ is the mass-thickness through the objects area and μ_i depends on the angle of acceptance of the aperture α and the scattering cross-section σ [16]. Thus, it can be seen that the number of electrons hitting a pixel depends on a monotonically increasing function of the mass-thickness. Therefore, the contrast in micrographs comes from the density variation in the specimen projected along the electron path. In particular, the natural logarithm of the number of electrons per pixel would consist of a projection of the specimen's density [16]. In addition, the large depth of focus in the TEM contributes to the projection requirement [17, 18]. This requirement holds for sufficiently thin samples in Bright Field (BF) microscopy, where mostly single scattering can be assumed [15].

The second type of contrast is the phase contrast. This type of contrast comes from the interaction of electrons with the electrostatic potential of the specimen, which can be represented as a "multiplicative specimen transparency function", explained in detail in [16]. This interaction will result in different phase shifts between the scattered and unscattered electrons. It is represented as a function that multiplies the Fourier coefficients of the wave, or equivalently, a transfer function that convolves the projection of the specimens electrostatic potential [18], known as the Contrast-Transfer Function (CTF). The CTF modulates the spatial frequencies and depends on the imaging conditions (acceleration voltage and defocus). Basically, if these conditions hold, the image can be considered as a projection of the electrostatic potential of the sample. Furthermore, as Hawkes states [16], the thinner the sample and the higher the intensity of the electron beam (controlled by the acceleration voltage) the better the projection requirement applies, and the better images approximate to projections.

[2]The mass-thickness μ consists of the multiplication of the density ρ and thickness t of the specimen.

[3]Note that usually Bright Field (BF) images are negatives of the originals. The intensity has been inverted so dark spots become light and vice versa.

Different preparation of the samples requires different imaging techniques. For example, in stained samples the imaging is based on amplitude contrast, while on freeze-hydrated samples comes from phase contrast [18]. The images used in this study consist of BF TEM images of biological specimens, attaining the projection requirement [3]. However, it must be noted that in the case of material science applications BF TEM images does not always attain the projection requirement [23]. But not only BF images fulfill the projection requirement in electron microscopy, other imaging techniques attain this requirement, such as High Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF- STEM), also known as Z-contrast microscopy, and Energy Filtered Transmission Electron Microscopy (EFTEM) [3, 23]. However, according to Hohmann [19] these two methods do not always ensure low electron dose, thus not being free of artifacts due to high electron doses that can damage the specimen.

Consequently, provided that the conditions are sufficiently met, the obtained images constitute projections of the inner electrostatic potential, or density distribution of the specimen. In this manner the micrographs can be used to perform an image reconstruction based on projections and the obtained tomogram should be a faithful approximation of the original specimen.

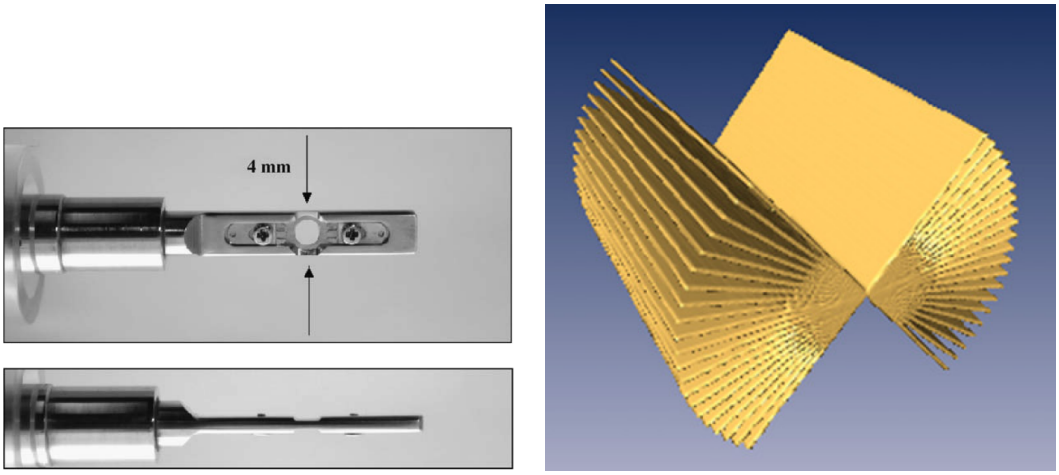
2.3 The Missing Wedge and Other Problems in Electron Tomography

Some problems are common in ET, being the most important, the missing wedge, the noise (low signal-to-noise ratio) and the incorrect alignment of projections. These problems are briefly described here in order to understand their effects on reconstruction better. The purpose of this study consists precisely of correcting one of them, the missing wedge.

2.3.1 The Missing Wedge

The principles of ET have been already described. A thin specimen is tilted over a fixed tilting axis and a set of projection images is acquired over a wide range of angles with a certain angular increment. Ideally, this should be performed over the complete angular range between -90° to 90° to obtain a complete reconstruction of the specimen, with as good angular sampling as possible. However this is not possible, the tilt range is limited, leading to a "wedge shaped" blind region that remains uncovered. This is known as the *missing wedge*, see Figure 2.6.

A complete angular imaging coverage of the specimen is impossible due to factors that impede acquisition at high angles. The design of the holders in the TEM, see Figure 2.6, is such that at high tilt angles, the holder becomes positioned between

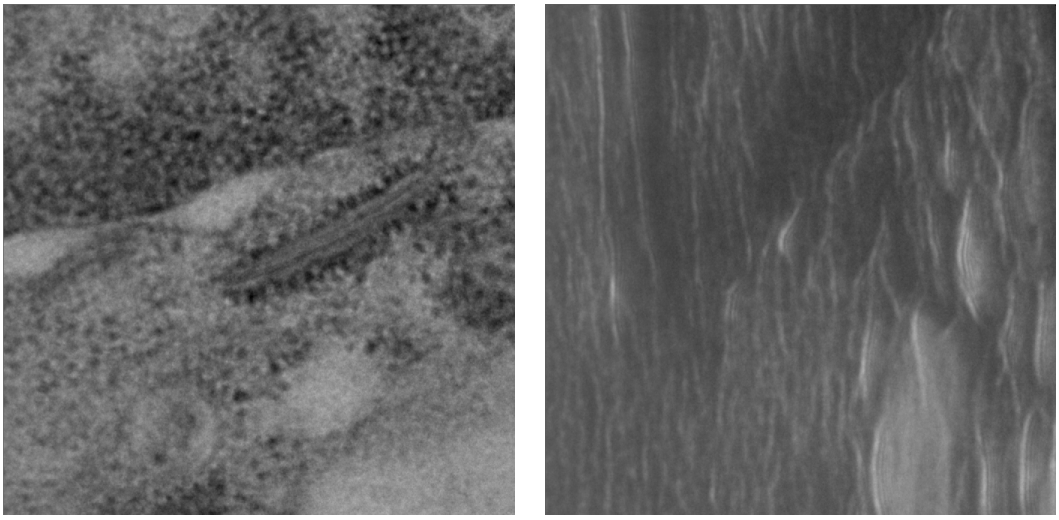


(a) Holder *(b) Missing wedge*

Figure 2.6: *(a) Example of a holder for a TEM specimen [24]. (b) Coverage for the single-axis tilting acquisition. A wedge shaped region remains uncovered [25].*

the electron beam and specimen shadowing the latter and making the acquisition impossible. The limited angle is not only determined by the holder, but also angles are restricted by the type and the position of the grid holding the specimen and the thickness of the specimen [26]. Other problems arise from the mechanical properties of the holder, mainly from its imperfect accuracy. The main problems include alterations in the ideal tilt geometry like imperfect eucentricity of the holder, instability and drift of the specimen during acquisition [5]. These lead to imperfections in the tilt angle and image warp, skew, rotation and magnification [26]. The mechanical imperfections in the goniometer will lead to displacements of the tilt axis of several hundreds of nanometers producing image shifts and even focus changes [17]. In addition, if imaging conditions are changed during imaging acquisition it can cause complications. According to Lucic *et al.* [18], it is important to keep certain requirements while recording a tilt series. The object of interest must remain within the field of view during the whole series, and all images in the series should be taken under similar imaging conditions.

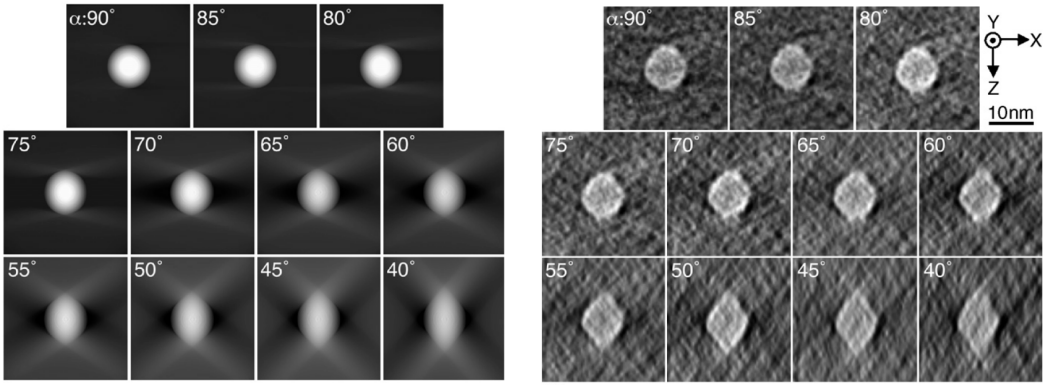
Not only the mechanical characteristics of the holder hinder complete angular range, but also it is affected by the scattering characteristics. Even though better holders were designed, still another major limitation for imaging at high tilt angles would be the increase in the specimen thickness. The longer path the electrons have to traverse through the specimen, the higher the probability of multiple scattering is, inelastic scattering in particular. This makes the projection requirement weaker, resulting in a degraded image, as can be seen in Figure 2.7. As an example, Baumeister illustrates this problem [8] : a 200 nm thick specimen at 0° tilt will increase up to 540 nm at a 70° tilt, being thicker than the mean free path for inelastic scattering, that is around 200 nm, resulting in an appreciable image degradation,



(a) 0° tilt image *(b) 70° tilt image*

Figure 2.7: *At high tilt angles images become degraded due to multiple scattering and artifacts start to appear.*

known as chromatic aberration [16]. The problem of image degradation due to the inelastic scattering can be partly solved with an energy filter, by which the electrons, which have suffered energy loss (due to inelastic scattering), are filtered out. Hence, it is possible to have thicker specimens up to 500 nm or even 1 μm [8], but with an inferior reconstruction quality. However, this technique will not perform correctly for thin specimens where elastic scattering is predominant [16].



(a) Simulated results *(b) Real results*

Figure 2.8: *Missing wedge reconstruction effects on the x-z plane depending on the maximum tilt angle [27]. Maximum tilt angles are shown in the figure.*

This limitation in the maximum tilt angle leads to reconstruction imperfections due to the missing wedge. The lack of projections within the missing wedge causes undersampling in the Fourier space, consequently degrading the reconstructed object in this direction [23]. The strength of these artifacts are dependent on the maximum tilt angle, and the lower this angle is, the more severely the reconstruction will be

affected. A representation of this effect has been properly illustrated by Kawase *et al.* in [27] and is shown in Figure 2.8. The amount of missing information due to the missing wedge can be estimated. Concretely the fraction of missing information q in Fourier space depends on the maximum tilt angle $\theta_{\max}$ and the volume of the slice [28],

$$q = \frac{1}{2}\frac{L_z}{L}\cot\theta_{\max}, \tag{2.4}$$

where L_z is the thickness of the slice, L is the size and it is assumed that $L_z \leq L$ and $\theta_{\max} > 45°$. As mentioned, this missing information produces distortions and imperfections in the reconstructed images. Particularly, it produces an elongation of the image features in the direction of the missing wedge (the z axis) that depends directly on $\theta_{\max}$ and determined by the elongation factor e_{yz}, [7, 29],

$$e_{yz} = \sqrt{\frac{\theta_{\max} + \sin\theta_{\max}\cos\theta_{\max}}{\theta_{\max} - \sin\theta_{\max}\cos\theta_{\max}}}. \tag{2.5}$$

This elongation artifact can be seen clearly in the images in Figure 2.8, both in simulated results and the real reconstructions. The lower $\theta_{\max}$ in the acquisition is, the more elongated the features in the image appear.

(a) *(b)* *(c)*

Figure 2.9: *Effects of the missing wedge on a reconstructed image for a maximum tilt angle of 60° [28]. The structures, such as membranes, contained in the direction of the missing wedge are completely unresolved. (a) Original image (b) Reconstruction with missing wedge in x-axis direction (c) Missing wedge in y-axis direction.*

2.3.2 Resolution in Electron Tomography

The Crowther's criterion (or Crowther's condition) gives an estimate of the obtainable resolution r of a tomographic reconstruction of a specimen of thickness d in the beam direction [18],

$$r = \frac{\pi d}{N}, \tag{2.6}$$

where N is the number of different projections acquired in the tilting stage. However, this is only applicable within the known range of projections. The missing wedge causes anisotropic resolution in the reconstructed tomogram, therefore the resolution will not be the same in every direction. This results in elongation artifacts in the direction of the missing wedge due to undersampling, or non sampling at all. As a result, structures, such as membranes, planar in the x-y plane [4] will be barely resolved in that direction [18, 28], see Figure 2.9. The decrease of resolution in the missing wedge direction (z axis), r_z can be determined by the elongation factor e_{yz} presented in equation 2.5 [7],

$$r_z = r_y \cdot e_{yz}. \tag{2.7}$$

This equation directly states that the higher maximum tilt angles yields better resolution in z axis, leading to a more isotropic resolution in the overall reconstruction. According to Plitzko and Baumeister, the final resolution of the reconstructed object does not depend only on the Crowther's condition, but also on the noise and the structural preservation of the specimen [7].

2.3.3 Noise Propagation in Electron tomography

Noise in the reconstructed tomogram is an accommodation of all noise types in the micrographs in the tilt series, and noise propagation strongly depends on the used reconstruction method [20]. This makes characterization of the noised difficult in the reconstruction. Furthermore, the deterministic effects due to the missing wedge in combination with the low Signal to Noise Ratio (SNR) makes the reconstruction and the interpretation of the tomogram even more difficult. If the structures are repeated along the tomogram, and assuming additive noise, a good approach to improve the SNR would be to use averaging, as done in single-particle reconstruction. However, this is not the case when imaging unique and non-repetitive structures such as cells and other biological samples. So the usual approach is to use filtering *a posteriori* to improve the reconstructed tomogram, where different filtering techniques can be used and some of them are explained in [20].

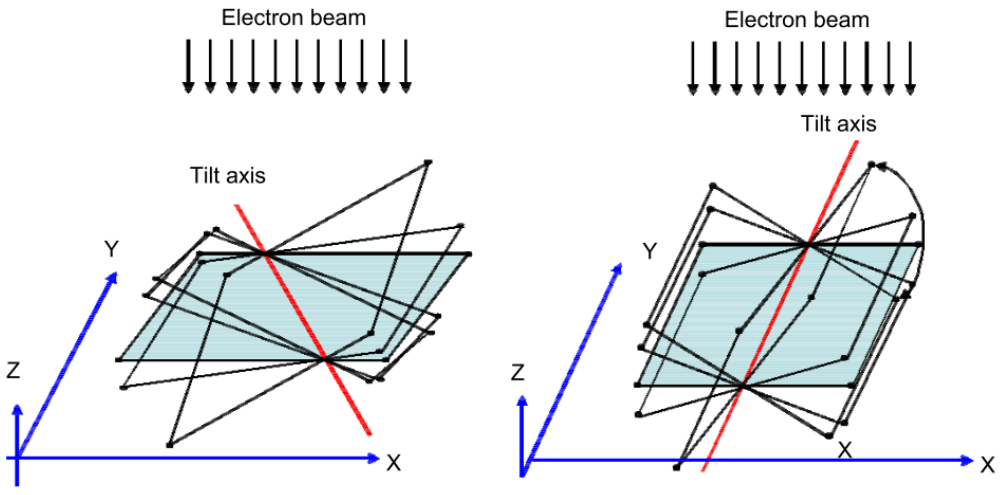
The noise distribution over the reconstructed set depends on the employed reconstruction methods and on the used imaging technique. However, noise-signal independence can be assumed in linear reconstructions such as Filtered Back Projection and Weighted Back Projection [20]. In addition, in the case of single-axis

[4] It refers to the structures perpendicular to the plane defined by the electron beam and the tilt axis [18].

tilting, noise variance is constant within the volume of isotropic resolution defined by the Crowther's condition [20], defined in equation 2.6. For ice embedded samples van der Heide *et al.* [30] found that the noise distribution in the reconstruction volume resembles quite well a Gaussian distribution, while for negatively stained samples the estimated noise distributions had heavier tails than Gaussian distribution. Furthermore, they detected that any reconstruction of stained samples seemed to have heavy tails in the noise distribution, characteristic of Poisson distribution.

2.3.4 Misalignment of Projections

An important problem present in ET is the alignment between consecutive projections and the location of the tilt axis in the TEM images. Due to the mechanical imperfections of the acquisition system, the obtained projections suffer some alterations with respect to the ideal projections. These include translational misalignment, rotational misalignment, image warping, skewness, rotation and magnification changes [26]. Automated electron tomography takes care of fixing these problems at every tilt of the specimen. Thus possible misalignments and change in focus area, in principle, are fixed during the acquisition. However, some of these imperfections still remain in the data set, so further processing is needed as a pre-processing step before reconstruction takes place.



(a) Rotated tilt axis *(b) Aligned tilt axis*

Figure 2.10: *Rotational offset of the tilt axis [31]. (a) The tilt axis is not parallel to the y-axis (b) The tilt axis is correctly adjusted*

The determination of the tilt axis in each image is also an important task. The position of the tilt axis should be parallel to one of the image axes and should be correctly placed in the middle position of the image axis perpendicular to the tilt axis, see Figure 2.10. Otherwise, reconstruction is compromised and strong

deterministic artifacts appear. Usually, tilt axis rotation is determined in the TEM microscope, and thus it can be corrected by simple image rotation. If the tilt axis position is not known, more complicated procedures exist for determining the tilt axis position, see for example [31]. If the axis is not correctly aligned with respect to the image axes, the center of rotation varies with one of the image axis (y-axis), and thus the artifacts will be more severe in the extremes, see Figure 2.11.

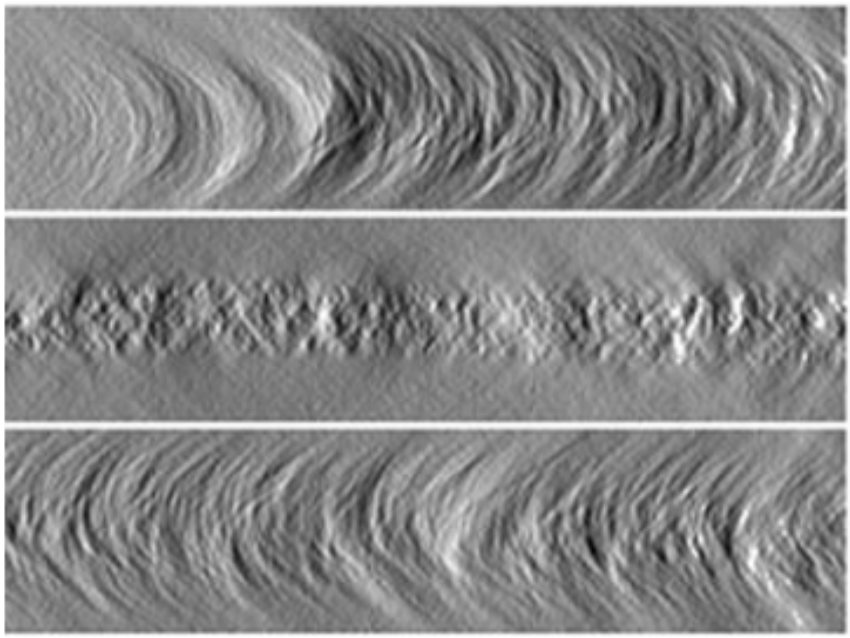


Figure 2.11: *Effects of the misaligned tilt axis [31]. Three different slices of the reconstruction in the top middle and bottom xz-plane are shown. The artifacts are more severe in the extreme slices due to the bigger error in the position of the center of rotation.*

3. IMAGE RECONSTRUCTION FROM PROJECTIONS

In fields such as medical imaging and scientific imaging it is desired to obtain images, or some estimation, of the internal structure of an object. This can be done by taking projections of the mentioned object at different angles and then using these projections to reconstruct the object's interior part. And such method is used, among others, in Computerized Tomography (CT), Positron Emission Tomography (PET), Magnetic Resonance Imaging, and also in the field that this work covers, Electron Tomography (ET) or in other words, tomographic imaging using electron microscopy. Although the problem faced in this thesis regards reconstructing 3D objects, the theory needed to understand can be explained in a simpler manner with a 2D case, and easily expanded to 3D case. In this chapter, first the mathematics needed to understand image reconstruction from projections are explained and followed by some of the most used reconstruction methods.

3.1 The Radon Transform

The mathematical principle, on which reconstruction from projections is based, is the Radon transform. It establishes the basis for all tomographic reconstruction methods, by which an object can be completely reconstructed from an infinite set of line integrals (projections) over an infinite set of views at different angles. This does not contradict the underlying principle of the ET reconstruction stated in the work of De Rosier and Klug work [2], which is based on Fourier-slice theorem. As a matter of fact, reconstruction based on Radon transform and Fourier-slice theorem are closely related and represent the same concept as it will be proven later.

Firstly, in order to understand the Radon transform, the concepts of projection and line integral need to be explained. Given a finite 2D object, defined by some function $f(x, y)$, Jiri [32] defines the *ray integral* along the line L at a given angle θ_0 and a distance ρ_0 from a given coordinate system as

$$g_{\theta_0}(\rho_0) = \int_L f(x, y)dr. \tag{3.1}$$

It collects the contribution of each point of $f(x, y)$ along the line integral at angle θ_0 and distance ρ_0 into a single value. Thus, a combination of these line integrals for

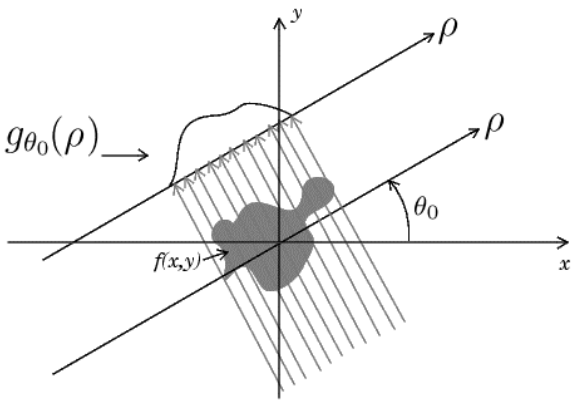


Figure 3.1: *Projection $g_{\theta_0}(\rho)$of an object $f(x,y)$ at an angle θ_0. Modified from [33]*

a given angle represents the *projection* of the function for that angle, see Figure 3.1. In reality, projection vectors correspond to physical phenomena such as, for example, the amount of attenuation of a beam of parallel X rays along its trajectory from the source to the detector through a sample.

If a set of projections at several angles is collected a *sinogram* is obtained,

$$g(\rho,\theta) = \int_{L_{\rho,\theta}} f(x,y)dr, \rho \in [0,\infty), \theta \in [0,\pi). \tag{3.2}$$

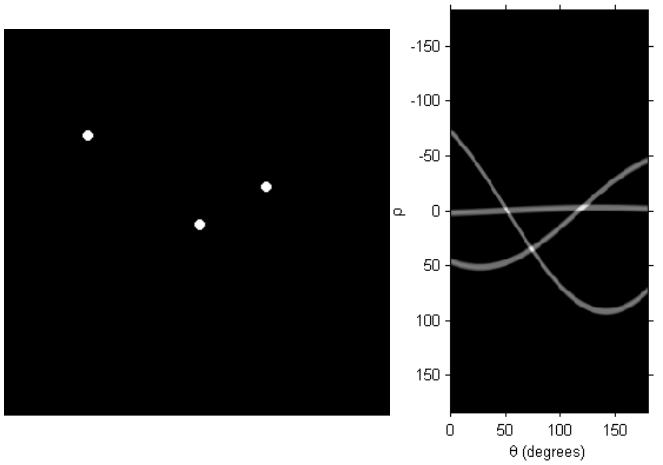
The sinogram $g(\rho,\theta)$ is also known as the *Radon Transform* of the function $f(x,y)$, and it is the basis for reconstruction of images from its projections. It constitutes a transformation from real space into the Radon space, where each point in the former space $(x,y) = (r\cos\phi, r\sin\phi)$ is a sinusoidal line in the latter, $\rho = r\cos(\theta-\phi)$, as can be seen in Figure 3.2.

A complete reconstruction requires infinite line integrals over infinite angles, however, such an approach becomes impossible in real situations. Thus, a more practical procedure uses finite summations to obtain an approximation of the original object or function. In addition, angular sampling is performed, so a finite set of projections is obtained. In this manner, the integration in equation 3.2 becomes a summation, yielding a discrete sinogram, defined in [34] as,

$$g(\rho,\theta) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\delta(x\cos(\theta)+y\sin(\theta)-\rho), \tag{3.3}$$

where x,y,θ and ρ are discrete variables and δ represents the discrete impulse function,

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$



(a) *$f(x, y)$ Original image* ***(b)*** *$g(\rho, \theta)$ Sinogram of (a)*

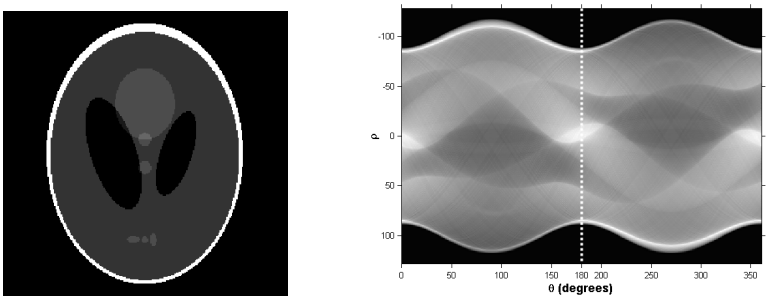
Figure 3.2: *Discrete radon transform. Each point in (a) creates a sinusoidal trajectory along θ in (b).*

Now, due to the sampling of the Radon transform, the reconstruction becomes an ill-posed problem, and no exact inversion can be derived. Then, the reconstructed object will be an approximation of the original, and for that reason many different methods have been developed.

A direct property of Radon transform is that any sinogram will be symmetric and periodic over θ with a periodicity of 2π,

$$g(-\rho, \theta \pm \pi) = g(\rho, \theta \pm 2\pi) = g(\rho, \theta). \tag{3.5}$$

Expressed in plain words, a sinogram has a periodicity of 2π radians and, in addition, it is symmetric around 0 (π) radians. Consequently, having a complete range of projections $[0, 2\pi)$ is redundant since the range $[\pi, 2\pi)$ corresponds to the mirrored version over ρ of the range $[0, \pi)$, see Figure 3.3. Therefore, from now on, it will be referred as a complete sinogram if the range $[0, \pi)$ ($[0°, 180°)$) is covered. However, although this property is present in Electron Tomography (ET) or X-ray Computed Tomography (CT), this is not always true in some real cases as it happens in the case of Single-Photon Emission Tomograpy (SPECT). In SPECT, a complete range of $[0°, 360°)$ is needed.



(a) Original image *(b) Sinogram of (a)*

Figure 3.3: *Periodicity and symmetry of the sinogram. In (b) the sinogram of (a) is periodic and symmetric, so a representation in the range* $[0°, 360°)$ *becomes redundant, and* $[0°, 180°)$ *is enough. The dotted line represents the axis of symmetry.*

3.2 Fourier-Slice Theorem

A fundamental relationship between the Radon Transform of an image and its 2D Fourier transform exists, known as the Fourier-slice theorem or projection-slice theorem. It states that the 1D Fourier transform of a projection of function $f(x, y)$ at a given angle corresponds to a slice that crosses the origin of the 2D Fourier transform of the same object at the same angle, as shown in Figure 3.4. Mathematically, it can be expressed as [34],

$$G(\omega, \theta) = [F(u, v)]_{u=\omega\cos\theta; v=\omega\sin\theta} = F(\omega\cos\theta, \omega\sin\theta), \tag{3.6}$$

where $F(u, v)$ is the 2D Fourier transform of the function $f(x, y)$,

$$F(u, v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)e^{-\jmath 2\pi(ux+vy)}dxdy, \tag{3.7}$$

and $G(\omega, \theta)$ is the 1D Fourier transform of $g(\omega, \theta)$ with respect to ρ

$$G(\omega, \theta) = \int_{-\infty}^{\infty} g(\rho, \theta)e^{-\jmath 2\pi\omega\rho}d\rho. \tag{3.8}$$

Furthermore, this relationship can be expanded for the 3D case, where a 2D Fourier transform of a projection plane of a 3D structure at a given tilt angle equals to a plane crossing the central section of the 3D Fourier transform of the object at the same angle. Thus, each projection contains information of a concrete part of Fourier space, so sampling it adequately for a complete range of angles it is possible to reconstruct the original volume. This is done by taking a collection of 2D projections and assembling their Fourier transforms in the 3D Fourier space, obtaining the original volume by just taking an inverse 3D Fourier transform [3, 19]. This is known as Direct Fourier Reconstruction (DFR) and it was the reconstruction method employed for the first tomographic reconstruction based on electron micrographs [3,

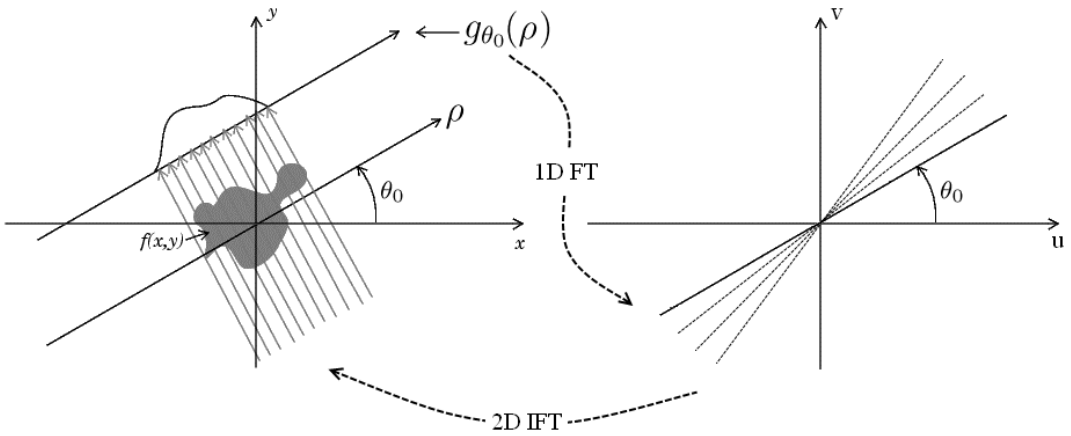


Figure 3.4: *Fourier-slice theorem in 2D. The 1D FT of a projection at an angle θ_0 corresponds to a radial line in Fourier space $F(u, v)$ crossing the origin at the same angle.*

see 9]. This method is explained in detail in this chapter.

Limited-angle tomography and the missing wedge

Following the Fourier-Slice theorem, the missing wedge problem can be easily explained. If the angular range during acquisition is limited, a region in Fourier space remains unsampled. This region is determined by the highest tilt angle, and in the 3D Fourier space, this uncovered region has a shape resembling a wedge. Hence, it is known as the missing wedge.

3.3 Reconstruction Methods

The basics of the Radon transform have been presented, and now the goal is to reconstruct the unknown object based on the given projections, or as stated by Carazo *et al.* "given a collection of projection images (2D data) g, determine the 3D structure f that produced the images g" [35]. Originally direct Fourier reconstruction has been used in Electron Tomography reconstruction, but it has been substituted by other methods. Nowadays, the most frequently used methods are Filtered Back-Projection (FBP) and Weighted Back-Projection (WBP) due to their simplicity and efficiency [23]. Also Algebraic Reconstruction Techniques (ART) and its variants are widely used in Electron Tomography.

3.3.1 Direct Fourier Reconstruction (DFR)

The first tomographic reconstruction of a TEM specimen, proposed by De Rosier and Klug [2], was based on the Fourier-slice theorem. By using a projection, the shape of an object can be partially described with the consisting frequencies [23]. Given this principle and collecting a series of projections at enough different views in a complete

range $[0°, 180°)$ ($[-90°, 90°)$ equivalently), the Fourier space becomes sampled. In this way, it is possible to fill the Fourier space of an object or distribution, making possible to recover it by means of a simple inverse Fourier transformation. This is known as Direct Fourier Reconstruction (DFR).

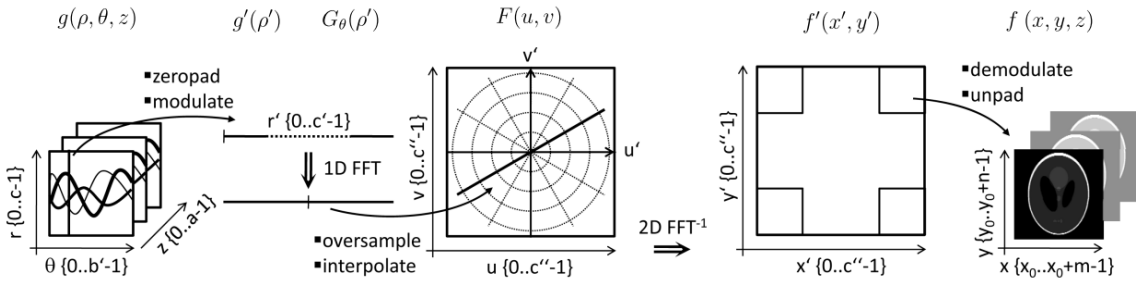


Figure 3.5: *Direct Fourier reconstruction [36].*

Using Equation 3.6, and representing the Fourier transform in polar coordinates $F(\omega, \theta)$, it is straight-forward to see that $F(\omega, \theta)$ is only known in some discrete points $\{\omega_i, \theta_j\}$, at the finite projection angles θ_j and sampled into discrete points ω_i. The problem arises when establishing a relationship between the two coordinate systems in order to construct the Fourier space in the rectangular Cartesian coordinates $F(u, v)$. Thus it is possible obtain the original object by simple inverse 2D Fourier transform [37]. This problem has been analyzed in detail and solutions to the interpolation problem have been provided [36–39]. A direct Fourier reconstruction method using interpolation in Fourier space was proposed in [36] and proven to give acceptable results. It uses the Fourier-slice theorem to build a polar 2D Fourier space of the object and maps it into rectangular Cartesian coordinates. After that, by means of inverse 2D Fourier transform, the original object is recovered, see Figure 3.5. The problem is solved in $O(N^2 \log N)$ complexity, outperforming the usual $O(N^3)$ in Filtered Back-Projection methods [36]. It uses zero-padding, oversampling in Fourier space and radial interpolation to overcome artifacts caused by the inadequate interpolation in Fourier space. Angular interpolation is linear. The mentioned algorithm is detailed in the appendix A, and further details can be found in [36].

3.3.2 Filtered Back-Projection (FBP) and Weighted Back-Projection (WBP)

An intuitive approach to reconstruct the original object is to try to go "backwards" in the process that created the projections. This is known as back-projection. It consists of *smearing* all projections along the directions at which they are acquired [34]. In 2D reconstruction, this is equivalent to replicate all 1D projection vectors along the direction of the projection angle and sum them. 3D reconstructions are

analogous to the 2D reconstruction. It involves replicating 2D projections along the third dimension in the beam direction. In this manner, each projection contributes to the reconstruction, and therefore the more projections acquired, the better the reconstruction is. Mathematically, the back-projection operation at an angle θ is expressed in [34] as,

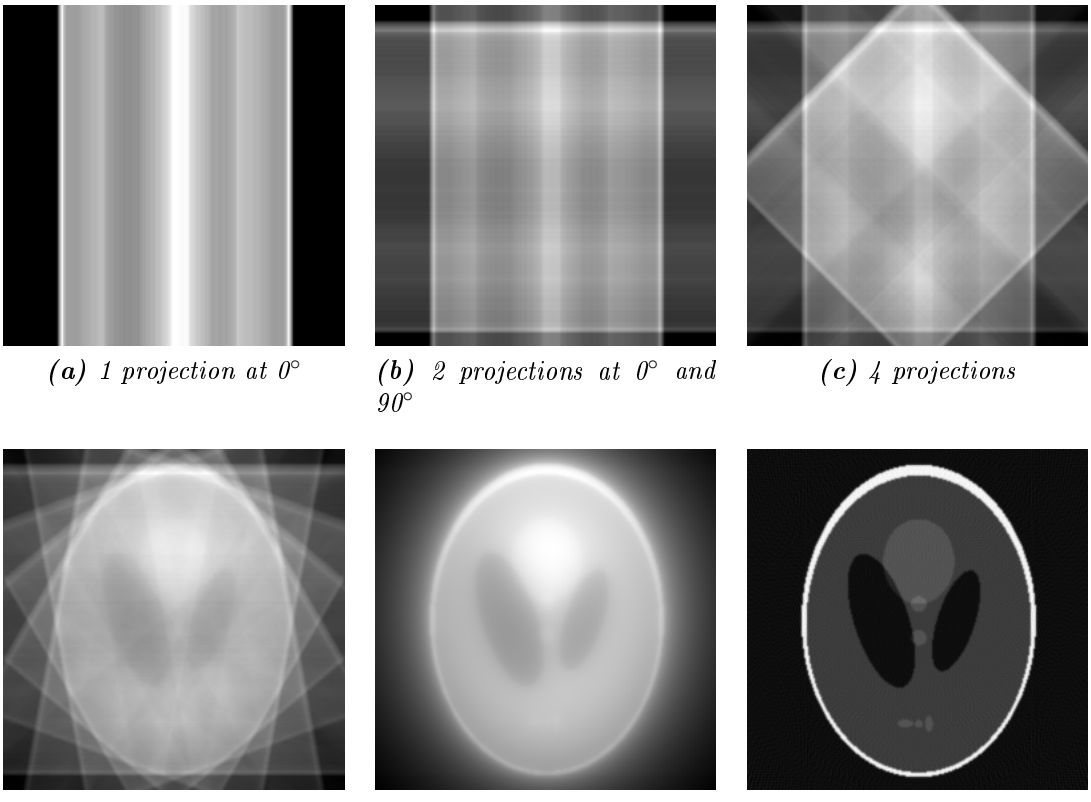
$$f_\theta(x, y) = g\left(x\cos\theta + y\sin\theta, \theta\right). \tag{3.9}$$

In order to reconstruct an image, back-projection should be performed by integrating for all projection angles from 0 to π,

$$f_{back-projected}(x, y) = \int_0^\pi f_\theta(x, y)d\theta. \tag{3.10}$$

For the discrete sinograms, the back-projected image is expressed as a finite sum of the discrete back-projections,

$$f_{back-projected}(x, y) = \sum_{\theta=0}^{\pi} f_\theta(x, y) = \sum_{\theta=0}^{\pi} g\left(x\cos\theta + y\sin\theta, \theta\right). \tag{3.11}$$



(a) *1 projection at 0°*

(b) *2 projections at 0° and 90°*

(c) *4 projections*

(d) *9 projections*

(e) *Back-projection of 180 projections*

(f) *FBP of 180 projections*

Figure 3.6: *Back-projection reconstructions. (a)-(e) Simple back-projection for different numbers of projections. (e) Filtered Back-Projection of 180 projections.*

This is known as simple back-projection, and different examples of this reconstruction for different number of projections are shown on Figure 3.6. However, this method creates a blurred image of the original object, see Figure 3.6, due to enhancement of the low frequencies caused by oversampling. Particularly, the obtained image corresponds to the 2D inverse Fourier transform of the Fourier transform of the object multiplied by a function of the form $\frac{1}{|\omega|}$ in the frequency domain [40]. A straight solution to this problem is filtering the projections before back-projecting [34], this can be deduced from the 2D Fourier inverse of the image Fourier transform,

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{j2\pi(ux+vy)}dudv. \tag{3.12}$$

Now, applying the Fourier slice theorem (equation 3.6) and a variable change, the equation 3.12 becomes

$$f(x,y) = \int_{0}^{\pi}\int_{-\infty}^{\infty} |\omega|\, G(\omega,\theta)e^{j2\pi\omega(x\cos\theta+y\sin\theta)}d\omega d\theta. \tag{3.13}$$

This is nothing but the back-projection presented in equation 3.10 of a filtered version of the projections, filtered by the 1D filter of the form $|\omega|$ in the frequency domain. This is known as the *Ram-Lak filter*, or just the *ramp filter*. This reconstruction is known as the Filtered Back-Projection (FBP) reconstruction and the result can be seen in Figure 3.6. In FBP reconstruction, the *ramp filter* is not always the best option as in the case of noisy reconstructions it will amplify high frequency noise. For this reason, different filters can be used, such as *Hanning* or *Hamming* windows, which attenuates high frequency noise.

A very similar reconstruction method is known as Weighted Back-Projection (WBP) where the projections are weighted by a weighting function that depends on the acquisition system geometry [41]. The acquisition system will determine the point-spread function, which is the optic equivalent to the impulse response in the field of signal processing. This point-spread function, denoted as $h(x,y)$ in 2D, will affect the output of the system by convolution [41],

$$g(x,y) = f(x,y) * h(x,y), \tag{3.14}$$

where $f(x,y)$ is the desired output of the imaging system, an image of a distribution in our case. As stated in [41], WBP consists of a back-projection followed by a deconvolution of the point spread function of the system. This is performed in Fourier space by dividing the Fourier transform of the distribution $F(u,v)$ to the Fourier transform of the point spread function $H(u,v)$. This is defined by a weighting function, which is expressed in 3D as,

$$W(u,v,w) = \frac{1}{H(u,v,w)}, \tag{3.15}$$

where u, v and w are the 3D coordinate axes in Fourier space. A general 3D form of the weighting function $W(u,v,w)$ in equation 3.15 will be defined by a set of acquisition angles θ_j and ϕ_j can be given in the form [41],

$$W(u,v,w) = \frac{1}{\sum_j 2a\text{sinc}\left[2a\pi\left(u\sin\theta_j\cos\phi_j + v\sin\theta_j\sin\phi_j + w\cos\theta_j\right)\right]}, \tag{3.16}$$

where a is the diameter of a sphere containing the sample, which for can be considered of infinite extent [41]. Now, if the single-axis tilting acquisition system is in question, the projections are taken over one axis at angles θ_j while $\phi_j = 0$. It can be assumed that $a \to \infty$ and $\theta \in$ [-90° +90°] [41], then,

$$W(\rho,\phi,v) = \rho, \tag{3.17}$$

where v is the tilting axis and ρ and ϕ are the cylindrical coordinates that replace u and w. Expressed in words, the weighting function W grows with the radius in the Fourier space ρ in the planes perpendicular to the tilting axis u, while it is constant along it. This radially linear function is a 2D version of the Ram-Lak ramp filter repeated along the tilting axis.

3.3.3 Algebraic Reconstruction Techniques (ART)

A different group of reconstruction methods are known as Algebraic Reconstruction Techniques (ART). These methods treat the object to be reconstructed as an array of unknowns and then create a set of algebraic equations for these unknowns based on the measured projections [42]. ART methods can be based as well on series expansions, where the reconstructed object is expressed as a linear combination of basis functions [35],

$$\hat{f}(r,\phi) = \sum_{j=1}^{J} x_j b_j(r,\phi). \tag{3.18}$$

The use of these basis functions allows the representation of the density function of the specimen in a general approach [43]. Usually, these basis functions consist of the voxels (or pixels) of the specimen. The approach is to estimate the unknown vector $\mathbf{x}$ based on the measurements $\mathbf{y}$. In this manner, if the measurements are assumed to be linear, the ith measurement (projection) is represented by y_i,

$$y_i \simeq \sum_{j=1}^{J} r_{i,j} x_j, \quad 1 \leq i \leq I, \tag{3.19}$$

being the values $r_{i,j}$ the corresponding element of the system matrix $\mathbf{R}$ that characterizes the data collection, or as described by Carazo *et al.* "the element $r_{i,j}$ is what the i-th measurement would be if the structure consisted of only the j-th basis function" [35]. The total number of measurements I is determined by the number and the size of the projections (*e.g.* for 120 projections of 512 pixels $I = 120 \times 512 = 61440$) and J by the number of pixels of the reconstructed image (or number of basis functions) with similar size. Since the system of equations create a huge matrix (*e.g.* 65000×65000), direct matrix inversion is not possible [42]. The ART method then solves the set of linear equations,

$$\mathbf{y} = \mathbf{R}\mathbf{x}, \tag{3.20}$$

by using an iterative approach. At each iteration the ith measurement y_i is used together with equation 3.19 to update the reconstruction obtaining a new estimation,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda \frac{y_{i_k} - \mathbf{r}_{i_k} \cdot \mathbf{x}^{(k)}}{\mathbf{r}_{i_k} \cdot \mathbf{r}_{i_k}} \mathbf{r}_{i_k}. \tag{3.21}$$

This means that at each iteration the ith measurement y_{i_k} is compared to the measurement that would give the current estimate $\mathbf{x}^{(k)}$, expressed by $\mathbf{r}_{i_k} \cdot \mathbf{x}^{(k)}$. This is, the current guess $\mathbf{x}^{(k)}$ is re-projected by using the projection operator $\mathbf{r}_{i_k}$ and compared to the measurement. The difference is then used to correct the actual estimate $\mathbf{x}^{(k)}$, using a step size λ that satisfies $0 < \lambda < 2$ [35]. This process is started with an initial estimation $\mathbf{x}^{(0)}$, usually set to zero, and then it is iterated for all measurements $1 \leq i \leq I$ for a given number of iterations. The solution of the ART method in the case of presence of noise is not unique, and in the case of undetermined system ($I < J$) infinite number of solutions may exist [42].

Regarding the size of the problem, considerable amount of computing power and storage are needed, so different versions of ART methods have been developed. A high performance ART approach that uses *blobs* as basis functions instead of voxels is presented in [35] and [43]. ART also allows the possibility of incorporating *a priori* information of the image in the reconstruction at each iteration, constraining the solution to be non-negative or zero outside a given area [42]. Other variations of ART are Simultaneous Iterative Reconstructions Technique (SIRT), where the update of the estimate is done after using all the measurements are used, and Simultaneous Algebraic Reconstruction Technique (SART) where the object is reconstructed in one iteration [42].

3.3.4 Maximum Likelihood Expectation Maximization (MLEM)

A different iterative approach is Maximum Likelihood Expectation Maximization (MLEM) reconstruction method [44–46], widely used for emission tomography. Iterations start from an initial guess of the image $\mathbf{x}^{(0)}$, which can be an FBP reconstruction of the projections (measured sinogram), or a uniform constant [45]. In each iteration, the current estimate is corrected by using the measured sinogram $\mathbf{y}$,

$$x_j^{(k+1)} = x_j^{(k)} \frac{1}{\sum_{i=1}^{I} r_{ij}} \sum_{i=1}^{I} \left[r_{ij} \frac{y_i}{\sum_{j=1}^{J} r_{ij} x_j^{(k)}} \right], \quad 1 \leq j \leq J. \tag{3.22}$$

The iterative updating is similar to ART, but using a multiplicative update instead of an additive one. At each iteration, the ratio between original measured sinogram and the re-projection of the current estimate creates a correction sinogram. The correction sinogram is then back-projected to the image domain, and this correction image is multiplied with the current reconstruction $\mathbf{x}^{(k)}$ to obtain the new estimated image $\mathbf{x}^{(k+1)}$. A flowchart of this procedure is shown in Figure 3.7. MLEM reconstruction retrieves an estimate of the specimen $\hat{\mathbf{x}}$, that maximizes the probability $p(\mathbf{y}|\mathbf{x})$ of observing the actual measurements $\mathbf{y}$ over all possible densities $\mathbf{x}$ [45]. As in ART, it is possible to incorporate known assumptions or *a priori* knowledge on the original density in the iterations.

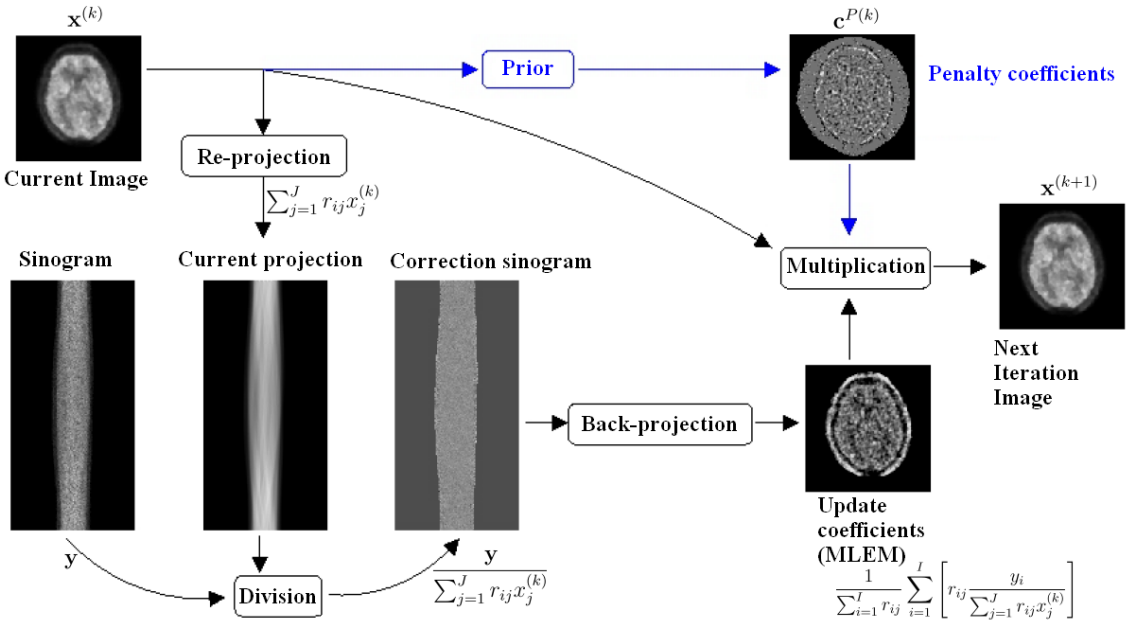


Figure 3.7: *Flowchart for an iteration for MLEM reconstruction [47]. Optional penalization in blue.*

Variations of the MLEM method exist. The Ordered Subsets Expectation Maximization (OSEM) method [48] divides the problem into subsets, by grouping the

projections into R ordered projection vectors. The reconstruction is divided into sub-iterations, and in each sub-iteration the update defined in equation 3.22 performed. This is done in the ordered manner, and the obtained estimation given by each subset is then used as the initialization for the next one. A complete iteration is achieved when all subsets have been used, and it can be repeated several times.

Other variations of the MLEM algorithm introduce penalizations in image domain at each iteration. The use of a penalty function can enhance the reconstruction in favor of some determined type of image by regularizing the reconstruction at each step. A common penalty function is the Median Root Prior (MRP) and its generalizations [49]. In standard MRP regularization, the image is assumed to be locally monotonic, and a median filter is used to eliminate outliers in the neighborhood of the estimated reconstruction. In the k-th iteration, a set of penalization coefficients $\mathbf{c}^{P(k)}$ is obtained from the penalty reference $\mathbf{A}$ to the current estimate, so it better meets the prior assumptions [49], see Figure 3.7. This penalization is then used to correct the current estimate,

$$x_j^{(k+1)} = x_j^{(k)} \frac{1}{\sum_{i=1}^{I} r_{ij}} \sum_{i=1}^{I} \left[r_{ij} \frac{y_i}{\sum_{j=1}^{J} r_{ij} x_j^{(k)}} \right] c_i^{P(k)}, \tag{3.23}$$

where

$$c_i^{P(k)} = \frac{1}{\sum_j r_{ij} + \beta \frac{x_i^{(k)} - A_i}{A_i}}. \tag{3.24}$$

In the case of MRP, the penalty A_i is the output of a median filter applied to the neighborhood of i. Note that in the literature other type of penalization exist [49].

4. GAP-FILLING APPROACHES

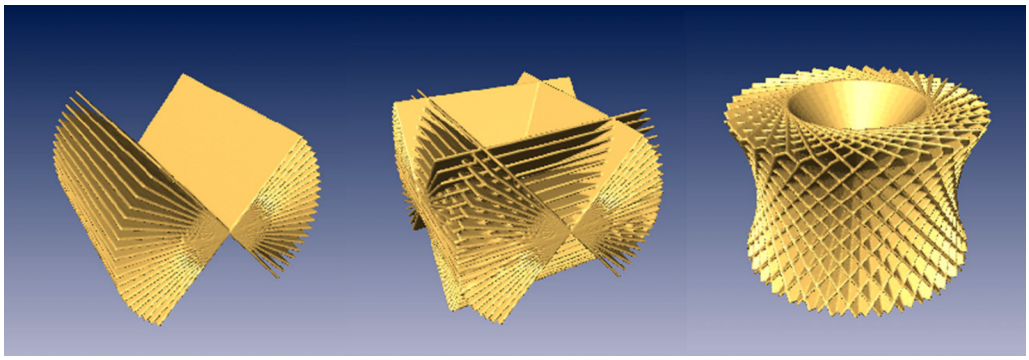
The main scope of this work is to perform a reconstruction of an Electron Tomography (ET) specimen after correcting the missing wedge present in this kind of imaging modality. In other words, it is desired to fill the gap in sinogram domain, or equivalently in the Fourier space, caused by the limited-angle acquisition scheme of ET. This chapter is divided in two separate sections. Firstly, a review of different gap-filling approaches or techniques to compensate the missing wedge is introduced. Secondly, a description of different attempts to solve the missing wedge problem is presented.

4.1 Current Approaches

Usually, reconstruction methods are not explicitly designed to deal with the missing wedge problem as in the case of Weighted Back-Projection (WBP). Although WBP is widely used because of computational simplicity and linearity, it is heavily affected by the missing wedge [43]. That is why, other solutions are provided in the literature. Some of them are briefly described in the following section.

4.1.1 Dual Axis and Conical Axis Acquisition

There are straight-forward solutions to overcome the missing wedge problem. One of the solutions is achieved by acquiring additional data from a perpendicular tilt axis [24]. This is known as dual axis tomography. Two series of images are collected tilting over two perpendicular axes. This technique effectively reduces the missing wedge into a *missing pyramid*, see Figure 4.1. This approach decreases considerably the missing information as can be seen in Table 4.1, where a comparison of covered Fourier coefficient in both methods is given. In particular, for double-axis tilt, the coverage of Fourier space is up to 93%, which is a more than acceptable percentage, although the missing information is mainly located in the missing pyramid direction, see Figure 4.1.



(a) Single axis *(b) Dual axis* *(c) Conical tilt*

Figure 4.1: *3D Fourier space representations of the three different acquisition methods. Single axis tilting creates a missing wedge, dual-axis tilting creates a missing pyramid and conical tilting creates a missing cone. [50]*

Tilt Range	Coverage	
	Single Tilt	Double tilt
±45°	50%	67%
±60°	67%	84%
±70°	78%	93%

Table 4.1: *Missing wedge and missing pyramid. Single axis tilting and dual axis tilting percentages of Fourier space coefficients covered [18].*

Nevertheless, dual axis tilting it is not always a correct solution since there is a more than 2-fold increase in total electron dose received by the specimen [1]. However, this can be solved by using sparse projections and taking care of redundancies by designing an appropriate angular distribution of projection angles [28]. In this manner, electron dose can be minimized while covering an acceptable percentage of the Fourier space.

More complicated acquisition techniques exist, such as conical tilting axis [50], where the missing wedge is reduced to a missing cone, as can be seen in Figure 4.1. However, conical tilting acquisition complicates the reconstruction process in electron tomography and it is mostly used in single particle reconstruction, see for example [25]. Finally, there exist proposed solutions in the acquisition system that even eliminate completely the missing wedge by using a rod-shaped sample is rotated over the complete range ±90° [27]. However, these solutions are not always possible, and thus the main focus is on correcting the missing wedge artifacts.

4.1.2 Use of *a priori* Information

Many methods present in the literature make use of constraints and *a priori* information to improve reconstruction and reduce artifacts in the limited-angle tomography.

In 1980 K.C. Tam and V. Perez-Mendez analyzed the effects of incomplete angle coverage and derived a method to recover the missing information [51], by using *a priori* knowledge of the finite extent of the object. Other methods may include, non-negativity, piecewise smoothness and amplitude constraints. These constraints are applied to the unknown parts of the reconstructed object or image while the known part is not modified. Some approaches include constraints on spatial support limits, maximum energy, density intervals and projections onto convex sets, but these solutions are not practical since they are useful only in some particular real cases [52].

In [53], a non-iterative linear extrapolator for the limited-angle tomography was proposed. It makes use of the bow-tie shaped spectral support of sinograms and sampling theory to extrapolate the missing part of the sinogram with the help of regularization. However, the proposed method needs a complete sinogram of a similar object to extrapolate the incomplete sinogram. Other method that use *a priori* information is defined in [54], where a Bayesian approach is presented. It consists of a Bayesian maximum *a posteriori* (MAP) reconstruction where it is possible to incorporate *a priori* probability densities, assuming that the object to be reconstructed belongs to an ensemble of similar images. More methods including *a priori* knowledge are reconstructions based on Ridge Functions [55] and based on projections onto convex sets (POCS) [56].

However, these mentioned methods are not straight-forwardly applicable to ET since the object to be imaged is often unique and unknown, making the application of strong constraints impossible. Constraints such as the limited extension, or limited volume, require precise knowledge on the specimen boundaries, or outer shape, attributes which are not known in general [18, 41]. As a consequence, the reconstruction methods based on prior assumptions, such as POCS, giving good results in ET are few [28]. Additionally, the introduction of wrong and strong assumptions leads to severe artifacts in the reconstruction [18]. Moreover, for the specimens embedded in negative stain, the surrounding stain affects the reconstruction if forced to be zero in a finite-extent constraint [41]. Also the complexity of the biological specimens does not allow to use similar sinograms for the reconstruction to use as *a priori* knowledge[5]. Furthermore, the misleading assumptions may cause incorrect structures and bias the results [41].

4.2 Organization of the Projections into Sinograms

Properties of sinograms are well known in biomedical image reconstruction, and a considerable amount of reconstruction methods and algorithms can be found based

[5]Note that this is not the case for single-particle reconstruction, in which precisely the repetitiveness of macro-molecules and other objects is exploited in the reconstruction.

on sinograms. In addition, many techniques have been developed for the limited-angle tomography based on sinograms, as shown in the previous section. These facts motivate the reconstruction of tomograms from the sinogram point of view, and therefore knowledge and methods in the field can be applied to the TEM projection images. Hence, as a first step, the acquired data (micrographs) should be re-ordered into sinograms. However, it is essential that the projections are perfectly aligned and centered, so as pre-processing step alignment should be performed. Alignment in ET is discussed in Appendix B.

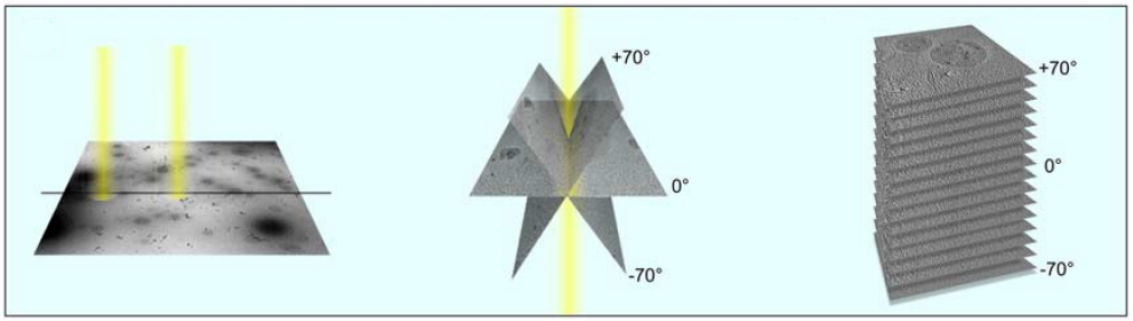
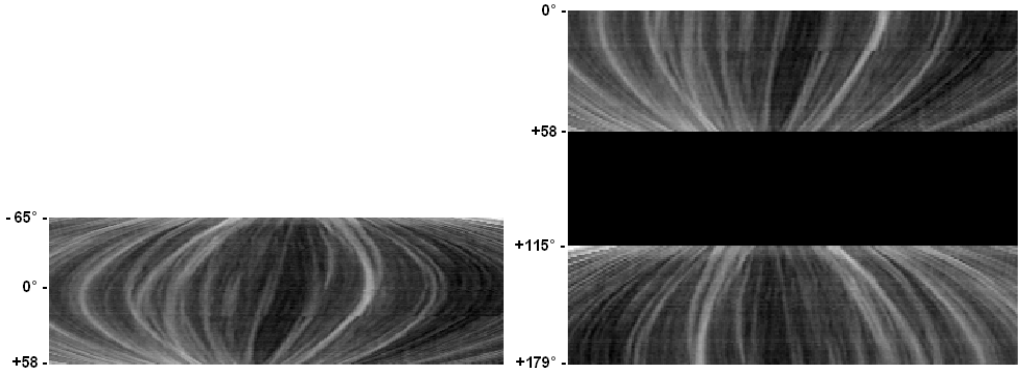


Figure 4.2: *Organization of pictures in ordered angle into a sinogram [26]. From left to right: Acquisition of a single TEM image. Acquisition of a tomographic tilt series of TEM images at different tilt angles in a range of* $[-70° + 70°]$.

As it has been described, a tomographic series of TEM images is acquired at different successive tilt angles. In addition, these images correspond to 2D projections of the sample. It is possible to arrange these images in angular order, for example from -70° to +70°, obtaining a 3D sinogram as can be seen in Figure 4.2. Moreover, sinograms are 2π-periodic and symmetric around 0 (π) and images in the range $\left[-\frac{\pi}{2}, 0\right)$, or equivalently in the range $[-90°, 0°)$, are the mirrored versions of the images in the range $[90°, 180°)$. Consequently, sinograms can be reordered into the range $[0°, 180°)$, and the gap corresponding to the missing wedge becomes evident, see Figure 4.3.

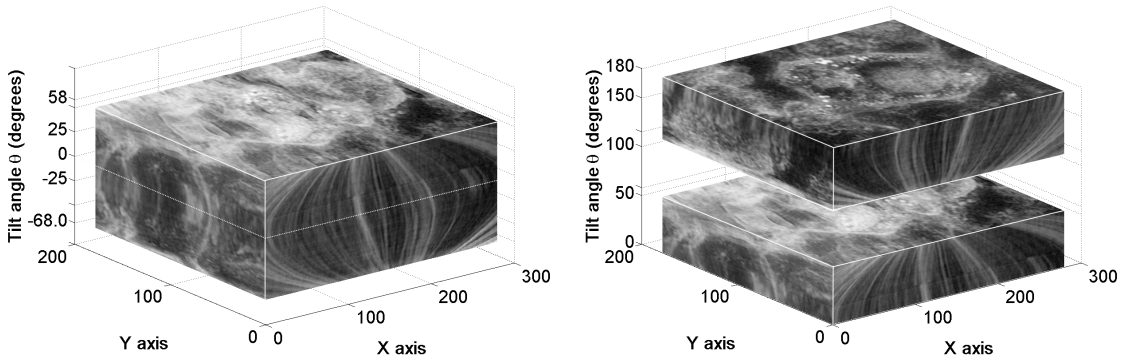
The obtained 3D stack, shown in Figure 4.4, is nothing but a stack of 2D sinograms along the transaxial direction (y-axis). This stack of 2D sinograms is used to divide the 3D reconstruction into a set of 2D reconstructions, simplifying the problem considerably and making it conveniently parallizable [57]. Each of the sinograms can be reconstructed separately, generating a stack of slices that can be joined together to form the complete reconstructed 3D image. This only holds if the stack of micrographs is correctly aligned and if the tilt axis is located parallel to the y axis in the 3D coordinate frame [57].

The division of the stack of the projection images into sinograms not only simplifies computationally the reconstruction, but it also aids the visual evaluation of the different reconstruction methods. After reordering the sinogram into the range



(a) Original sinogram, range $-65°$ to $+58°$ *(b) Range $0°$ to $+179°$*

Figure 4.3: *Due to sinogram periodicity and symmetry properties, projections at negative angles can be placed mirrored in a range from $90°$ to $+179°$.*



(a) Original 3D stack *(b) Re-ordered stack*

Figure 4.4: *Re-ordering of 3D stacks. The original stack (a), containing projection images in the range $[-65°, 58°]$, is re-ordered in the range $[0°, 180°)$ (b). Now the 3D gap corresponding to the missing wedge becomes evident. Note that each slice along the tilt axis (y-axis) is a 2D sinogram.*

$[0°, 180°)$, the missing wedge becomes explicit, in the middle of the sinogram. Since the aim of this work is to fill this gap, it is easier to evaluate visually the performance of the reconstruction method if the missing wedge is clearly visible. Furthermore, due to the Radon transform properties, any point in space will follow a sinusoidal trajectory in the sinogram, allowing a straight-forward pre-evaluation of the reconstruction method applied. This is, any successful reconstruction method should generate continuity of these sinusoidal trajectories within the filled gap.

4.3 Initial Approaches

In order to understand the missing wedge problem better and its effects on reconstruction, some initial attempts were performed. The first approach is to use known reconstruction methods on some known dataset without explicitly compensating the missing wedge and thus evaluate how it influences the result. After initial reconstructions, different procedures were tested, explained in the following subsections.

4.3.1 Iterative Projection - Back-projection

This method is based on the Radon-transform-iteration scheme proposed by Tom and Perez-Mendez in [51, 58]. The original method proposes an iterative method moving between projected space $g(\rho, \theta)$, and real space $f(x, y)$. This is performed iteratively by projecting and back-projecting, and using as a constraint the finite extension of the object in image domain. At each iteration, the finite extension of the object $f_k(x, y)$ is used as a constraint, by setting the values outside the known object position to zeros. The second constraint used is the known projections from the limited angular range in the sinogram. The block diagram of this procedure can be seen in Figure 4.5.

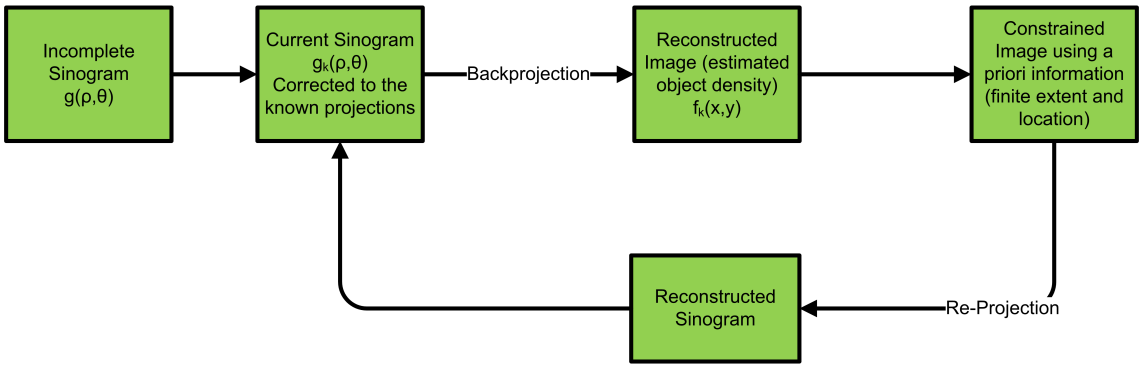


Figure 4.5: *Radon-transform-iteration scheme proposed in [51] and [58].*

However, the finite extension constraint cannot be applied in ET because *a priori* information is not applicable. Thus, a modified version was proposed with looser constraints and additional filtering in the image domain. The first constraint is the non-negativity constraint, as the image cannot have any negative values and consequently any negative pixel value is set to zero. Filtering in image domain is also used as a regularization. At each iteration, the current sinogram, $g_k(\rho, \theta)$, is back-projected to the image domain where constrains and filtering are applied. The resulting image, $f_k(x, y)$, is re-projected into the sinogram domain, and the information within the unknown gap is masked and inserted into the original sinogram, $g(\rho, \theta)$, Figure 4.6.

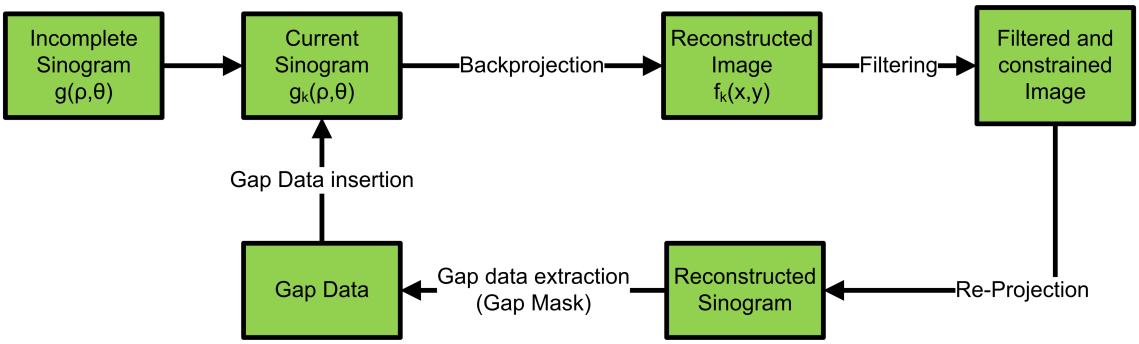


Figure 4.6: *Iterative Projection - Back-projection method.*

Other simple constraints could be inserted, such as energy conservation property of the Radon transform. However, for simplicity they are not included in the simulations.

4.3.2 Gap-Filling Using a Dedicated DCT-Domain Filter

Other type of methods tested were focused on a variation of the iterative procedure proposed by Tuna *et al.* in [59] and [60]. Their method consists of similar technique to the Constrained Fourier Space (CFS) method [59, see 4] and registered to outperform it. The method, originally proposed in [60] consists of a dedicated 2D Discrete Cosine Transform (DCT) domain filter. Two main reasons motivated the used of DCT instead of Discrete Fourier Transform (DFT). Firstly, DCT has higher energy compaction for highly correlated data. Secondly, the DCT coefficients corresponding to the gap have lower values compared to the DFT coefficients.

The method exploits the known shape of the gaps in sinogram domain of the High-Resolution PET sinograms to characterize its gap-dependent DCT coefficients. Different DCT coefficients are segmented using different thresholding levels, and then combined in a weighting mask that will act as a dedicated filter, see Figure 4.7. This experimental mask is iteratively used to suppress the DCT coefficients created by the gaps, effectively filling the gaps in the sinogram domain. The diagram of the procedure can be seen in Figure 4.8.

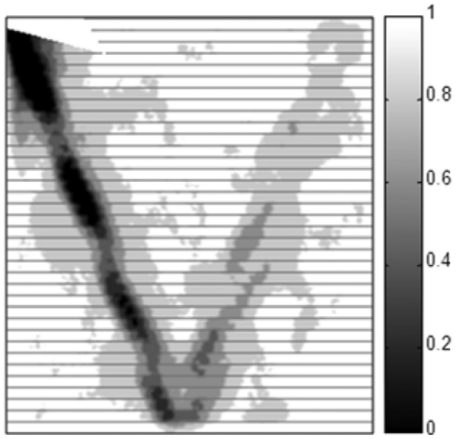


Figure 4.7: *Dedicated DCT domain filter for the HRRT PET scanner [59].*

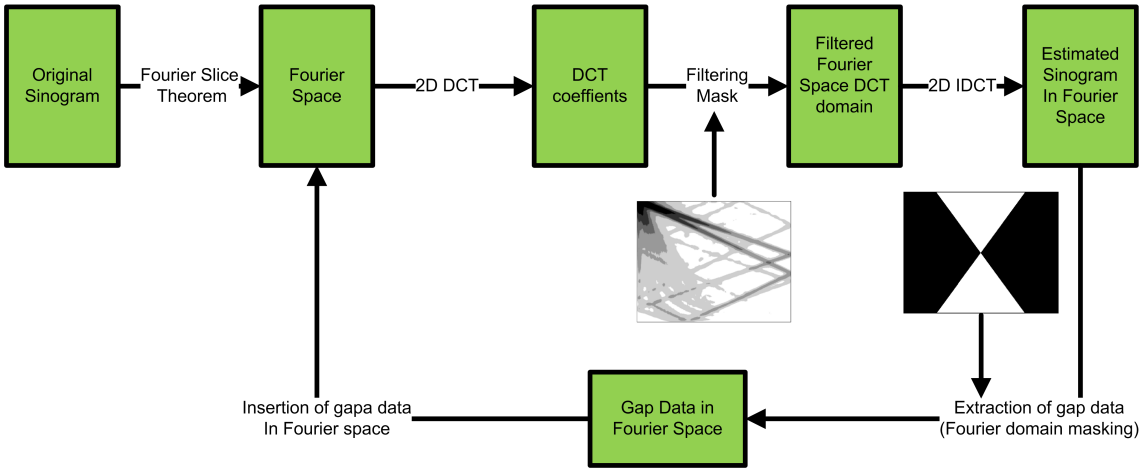


Figure 4.8: *Gap-filling using a dedicated DCT-domain filter.*

In this work, a variation of this dedicated DCT-domain filter was tested. As mentioned, the method exploited the known characteristic shape of the gaps to create a dedicated mask to filter out the DCT coefficients corresponding to the gaps. In the limited angle case in ET, the gap corresponding to the missing wedge has known characteristic shape in the sinogram, as shown in Figure 4.9. However, this characteristic gap shape is constant along the radial dimension, suffering changes only in the angular dimension, and thus does not have good properties in the DCT domain, concentrating the energy mostly in one dimensional line of DCT coefficients, see Figure 4.9. This is why another representation of the missing wedge was chosen. An equivalent form of displaying the information acquired in the projections is by representing them in the Fourier space. Given the Fourier-slice theorem, the gap has the characteristic wedge shape in Fourier space as shown in Figure 4.9.

The method consist then in assembling the Fourier space based on the known projections. The known angular coverage determines the shape of the missing wedge.

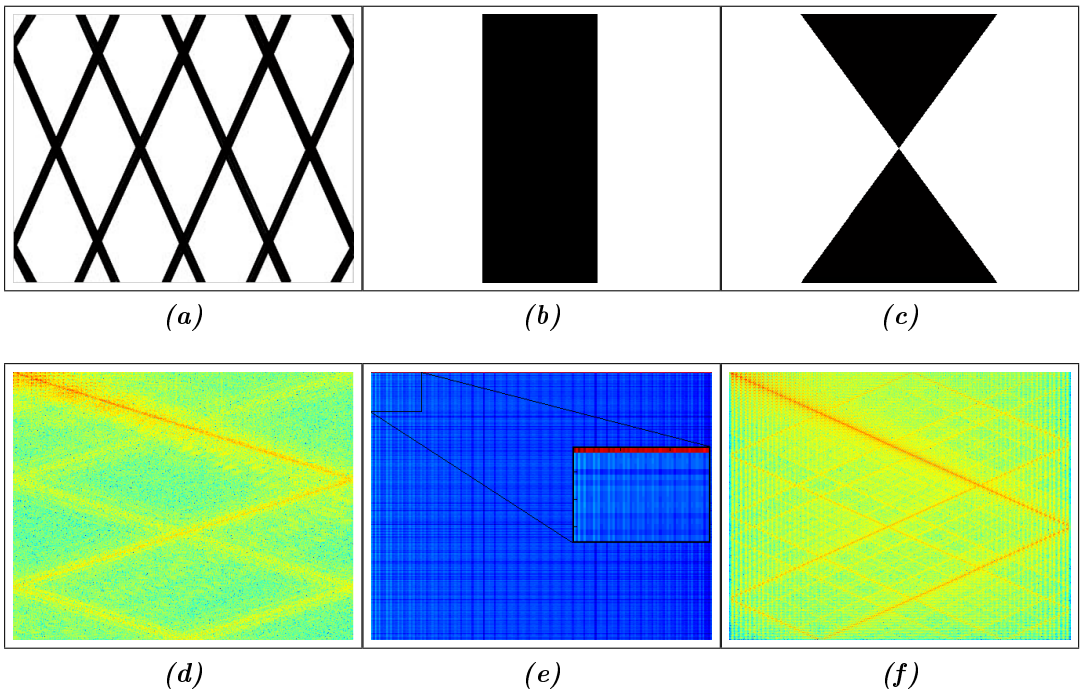


Figure 4.9: *Representation of gaps and their characteristic DCT coefficients. (a) Gaps in sinogram domain of HRRT scanner [59]. (b) Gap in sinogram domain caused by the missing wedge. (c) Missing Wedge in Fourier domain. (d) DCT coefficients of (a). (e) DCT coefficients of (b). (f) DCT coefficients of (d).*

The gap has distinctive DCT coefficients, see Figure 4.9, which were used to create the dedicate mask as described in [60]. Thus, the higher values these coefficients have, the more they are attenuated and the effects of these coefficients are reduced. Compensation of the gaps in Fourier space was achieved. Then, after Fourier space gap has been filled by this method, the image can be obtained by simple inverse Fourier transformation as in direct Fourier reconstruction technique.

4.3.3 Image Inpainting

An interesting approach is to treat the sinogram as an image and apply the method know as image inpainting. The term inpainting comes from the techniques used by restorers and painters to modify, or recover, missing and damaged parts in a painting in an unnoticeable way. The method that replicates this technique for digital images was proposed by Bertalmio *et al.* in [61]. This method was already proposed for interpolating small gaps in the sinogram giving good results [62].

The image inpainting algorithm suggested in [61] uses a set of partial differential equations to propagate the curves arriving to the gap regions. In simple words, the method tries to fill the missing or damaged parts, called here gaps, by using the information surrounding these gaps. This is done by continuing the isophote

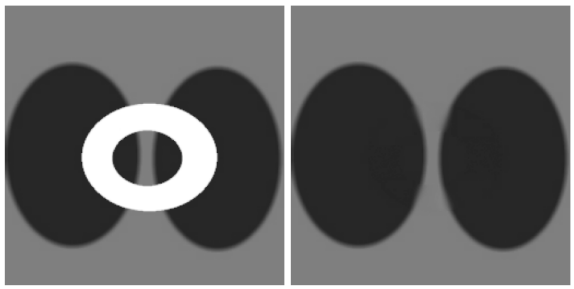


Figure 4.10: *Example on image inpainting with a synthetic figure [61].*

(curves containing the same light intensity level) arriving to the gap boundary. An example of the result of applying this method is shown in Figure 4.10. The process of inpainting is briefly described below.

At each iteration n, the gap Ω inside the image $I^n(i,j)$ is updated by using $I_t^n(i,j)$ with a rate of Δt,

$$I^{n+1}(i,j) = I^n(i,j) + \Delta t I_t^n(i,j), \quad \forall (i,j) \in \Omega, \tag{4.1}$$

where the update $I_t^n(i,j)$ is calculated using the information surrounding the gap Ω, and according described in [61] as,

$$I_t^n(i,j) = \overrightarrow{\delta L^n}(i,j) \cdot \overrightarrow{N}^n(i,j), \tag{4.2}$$

where $\overrightarrow{\delta L^n}(i,j)$ represents the measure of change in the information $L^n(i,j)$ and $\overrightarrow{N}^n(i,j)$ is the propagation direction. In this manner, it is wanted to propagate some information $L^n(i,j)$ in the direction of $\overrightarrow{N}^n(i,j)$. Therefore, both terms need to be calculated. For the information, Bertalmio *et al.* propose to use the Laplacian as a smoothness estimator

$$L^n(i,j) = I_{xx}^n(i,j) + I_{yy}^n(i,j). \tag{4.3}$$

The direction of the propagation is calculated with a time varying estimator of the isophotes direction field, given by the perpendicular of the gradient $\nabla^{\perp} I^n(i,j)$, which gives the direction of the minimal spatial change, the isophotes direction [61].

4.4 Compressed Sensing (CS) Approach

The term compressed sensing, or compressive sampling (CS), refers to the concept of recovering a signal (a digital image in this case) from only few limited measurements. Formally, it can be expressed as reconstructing a signal x in $\mathbb{R}^m$ by measuring n general linear functionals of x, $n \ll m$ [63]. The signal is assumed to be described sparsely with respect to othonormal basis, such as Fourier basis, for $\mathbb{R}^m$. Moreover, this principle holds generally for natural signals and measurements [63]. Thus, the

signal can be reconstructed from an incomplete part of the spectrum given a linear transform, using non-conventional reconstruction methods. Note that usual methods do not work, for example in the strongly ill-posed case of limited angle tomographic reconstruction [63, 64]. Concretely, limited angle tomographic reconstruction is the problem faced in this work. In this manner, the mentioned linear functions are in particular Radon projections, and thus given the central-slice theorem, they can be represented sparsely with the Fourier transform, as can be seen in Figure 4.11.

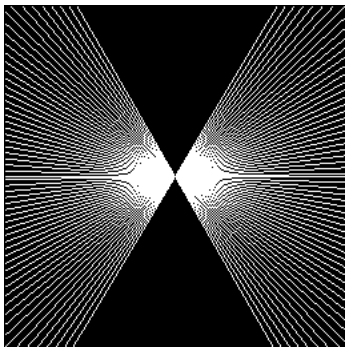


Figure 4.11: *Projection operator S representation in Fourier space. Given the Fourier slice theorem, each line corresponds to a projection of the object at the same angle.*

4.4.1 Compressed Sensing Algorithm

In their work "Compressed Sensing Image Reconstruction Via Recursive Spatially Adaptive Filtering" Egiazarian *et al.* [64] present a novel approach for the inverse problem in tomographic reconstruction, inversion from sparse projections and limited angle tomography. Furthermore, they attain exact reconstruction in these two cases for the synthetic data.

The method consists of a non-parametric stochastic approximation regularized by an adaptive filter. At each iteration, "the algorithm is excited by injection of random noise in the unobserved portion of the spectrum and a spatially adaptive image denoising filter, working in the image domain, is exploited to attenuate the noise and reveal new features and details out of the incomplete and degraded observations." [64]. The algorithm characterizing the recursive system is described in Figure 4.12. where the spectrum of the image given by the 2D transform T is represented by y. The transform T is the 2D Fourier transform, and the spectrum is then divided into the known as portion $y_1 = S \cdot y$ and unknown part as $y_2 = (1 - S) \cdot y$ by the sampling operator or binary mask S, which tells exactly where coefficients of the spectrum are available, see Figure 4.11.

The algorithm can be described formally by a recursion, where the k-th iteration can be expressed as,

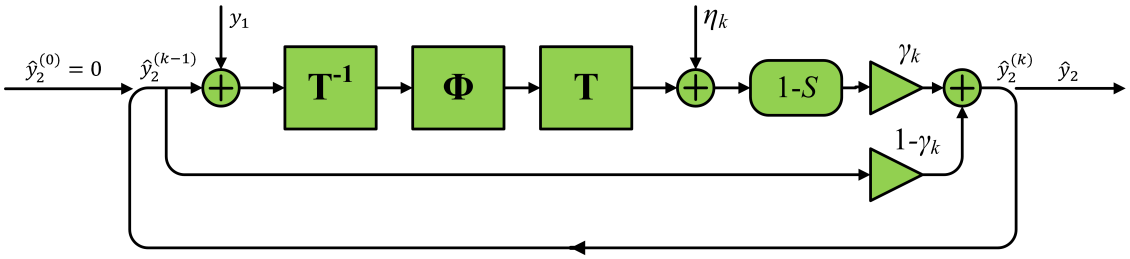


Figure 4.12: *Flowchart of the recursive CS procedure described in [64]. y_1 is the known part of the spectrum while the unknown part is initialized to zero, $\hat{y}_2^{(0)} = 0$.*

$$\begin{cases} \hat{y}_2^{(0)} = 0, & k = 0, \\ \hat{y}_2^{(k)} = (1-\gamma_k)\,\hat{y}_2^{(k-1)} + \gamma_k\,(1-S)\left[T\left(\Phi\left(T^{-1}\left(y_1 + \hat{y}_2^{(k-1)}\right)\right)\right) + \eta_k\right], & k \geq 1. \end{cases} \tag{4.4}$$

In the k-th iteration, the estimate of the unknown part of the spectrum is then represented by $\hat{y}_2^{(k)}$, thus the current estimate of the complete spectrum is,

$$\hat{y}^{(k)} = y_1 + \hat{y}_2^{(k)}. \tag{4.5}$$

As expressed in equation 4.4, the unknown part of the spectrum is initialized by setting $\hat{y}_2^{(0)} = 0$. At each iteration, an estimate of the current image is obtained by inverse transform $T^{-1}\left(y_1 + \hat{y}_2^{(k-1)}\right)$ and is filtered by the spatially adaptive filter Φ. Following the filtering, random noise η_k is added to the unknown portion of the spectrum. The added noise will act as a random generator of the missing coefficients of the spectrum and the binary operator $(1 - S)$ will extract these coefficients and combine them with the previous estimation with a scale factor of γ_k, parameter that acts as a step size of the algorithm. [64]

This method can be treated as a stochastic approximation and the additive noise η_k acts as an accelerator of the process, and as the iterations continue the amount of added noise is decreased. This generates random walks around the global minima, in a similar way as in the simulated annealing [64]. The noise η_k should be high at the beginning, allowing the method to escape out of possible local minima and progress further to the global extremum as the algorithm iterates and the variance of the added noise decreases. In addition to this, the denoising filter should be strong enough and it should be possible to control the level of smoothing along with the variance of the added noise. According to [64], the step size γ_k can be fixed provided that the variance of η_k is decreased to 0 as the iterations progress to infinity.

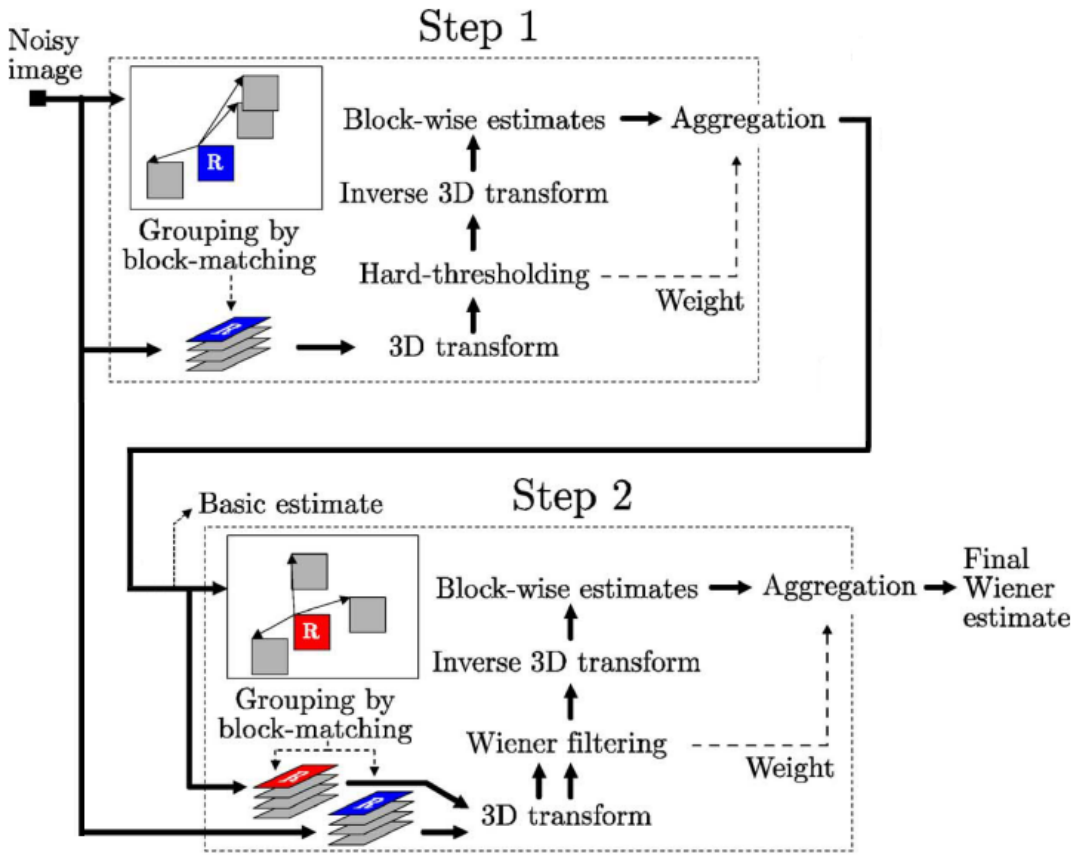


Figure 4.13: *Flowchart of BM3D filter [66].*

BM3D Filter

The CS method described in [64] uses a special spatially adaptive filter Φ which is the Block-Matching and 3D filtering algorithm (BM3D) presented in [65] and [66]. This filter takes advantage of the similarities between the parts present in images to create groups of blocks. Using a block matching procedure, 3D groups of 2D blocks are created and further filtered in a collaborative way. Each block in the groups helps filtering the other blocks. This is performed by exploiting the enhanced sparsity of these blocks in some transform domain to denoise effectively the image by thresholding the coefficients of the transform. A description of the BM3D filter is shown in Figure 4.13.

This BM3D filter is optimized for denoising additive Gaussian noise, and it is one of the best methods for denoising independent additive Gaussian noise present in literature. Following the assumptions about additive independent Gaussian noise present in ET [21], the BM3D filter results in a good filter for TEM images. Firstly, its strong denosising properties allow the reconstruction as described in the CS method. Secondly, it is a perfect selection for denoising the noise present in the TEM images, assuming that additive Gaussian noise is predominant.

4.5 Variations of the Binary Mask in the Compressed Sensing Approach

The reconstructions by using the CS algorithm performed in [64] can be interpreted as a type of Direct Fourier Reconstruction since it uses a set of known Fourier coefficients in order to obtain the original image. In this manner, it must be noted that it is not a sinogram based reconstruction. In their approach, a set of Fourier coefficients are segmented from the complete spectrum, simulating the limited angle Radon projections by taking radial lines in the Fourier space by using the sampling operator S. This is used as the known part of the spectrum y_1.

However, this is not straight-forward when dealing with imperfect discrete projections as occurs in the ET, where both projection directions and gray level information are corrupted [35], the source is an incomplete discrete sinogram made up of imperfect discrete projections, and as such, there is no exact inversion for this problem (discrete Radon).

Since the positions of true coefficients are not completely known. If it can be assumed that the estimated Fourier coefficients within the known range are good enough (*e.g.* using interpolation), a complete mask covering this range could be used, yielding a bow-tie shaped mask, see Figure 4.14.

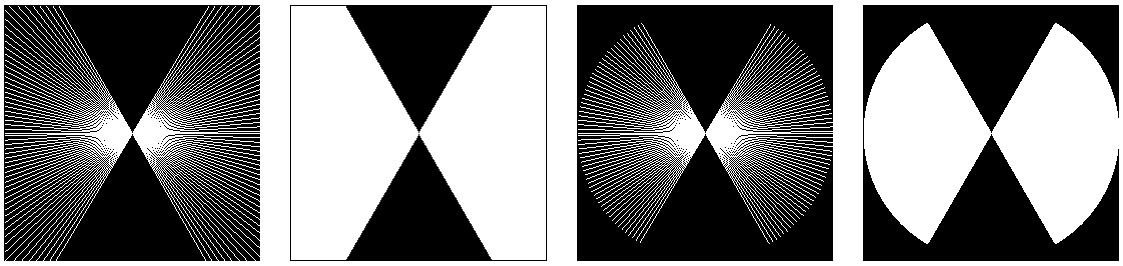


Figure 4.14: *Variations of the Fourier space binary mask S for a coverage of* $[-60°, +60°)$.

Assuming that the highest frequency components are mostly noise (as it certainly happens with real TEM data), this mask could be further modified to filter out the high frequency coefficients. Furthermore, if DFR is used, the Fourier space is calculated only within a circle of a radius determined by the length of the FT, forcing the high frequency Fourier coefficients outside this circle to be zero. In this case, the use of a filtered mask is justified. These reasons motivate the use of different masks in the simulations in order to evaluate the results achieved with each of them. The different proposed masks are shown in Figure 4.14.

4.6 Modifications of the Compressed Sensing Approach

The available data in ET are the measured projections, so a direct application of CS algorithm is not possible, needing a sinogram based approach. Given the central-slice theorem, it is possible to know theoretically how the projections in the sinogram relate to the coefficients of the Fourier space. However, this positions do not fall in exact pixel positions and therefore the coefficients that represent the known part of the spectrum, y_1, are not exactly known. If the wrong Fourier coefficients are segmented and used as the known part of spectrum, the reconstruction will present artifacts.

In this manner, a method to obtain y_1, or at least a good approximation, from the sinogram is needed. The main idea is to use some kind of initialization (or estimation) method M in order to obtain an initial estimate of the image $\hat{f}^M(x,y)$ and its spectrum $\hat{y}^M = \hat{y}_1 + \hat{y}_2^M$, where $\hat{y}_1$ is used as the known part of it and $\hat{y}_2^M$ corresponds to an initialized (estimated) unknown part of the spectrum. This estimate is then used as an initialization for the CS procedure, where the recursions are then defined as,

$$\begin{cases} \hat{y}_2^{(0)} = \hat{y}_2^M, & k = 0, \\ \hat{y}_2^{(k)} = (1-\gamma_k)\,\hat{y}_2^{(k-1)} + \gamma_k\,(1-S)\left[T\left(\Phi\left(T^{-1}\left(\hat{y}_1 + \hat{y}_2^{(k-1)}\right)\right)\right) + \eta_k\right], & k \geq 1. \end{cases} \tag{4.6}$$

The Fourier coefficients corresponding to the known part of the spectrum, now denoted as $\hat{y}_1$, are segmented from the initial estimate of the spectrum by the sampling operator S. Note that, ideally, $\hat{y}_1 = y_1$. The unknown coefficients $\hat{y}_2$ are initialized to $\hat{y}_2^{(0)} = \hat{y}_2^M$. Finally, these two portions of the spectrum are used as an initialization for the CS recursions to reconstruct the image. An estimate of the image can be obtained in the k-th iteration given by $T^{-1}(\hat{y}_1 + \hat{y}_2^{(k-1)})$, where the transform T is the Fourier transform (FT).

The motivation for using initialization lies on the fact that the CS method, as described in [64], consists of stochastic approximations excited by additive noise and driven by a spatially adaptive filter. For these approximations, the unknown part of the spectrum y_2 is set initially to zero, see Equation 4.4. Then, noise is recursively added to this unknown portion of the spectrum and filtered out by the adaptive filter in spatial domain. The excitation with noise helps to avoid possible local minima, converging to the global extrema [64]. If, by means of initialization of y_2, Equation 4.6, the initial estimate lies closer to the solution, the optimization should be closer to the global minima.

On the other hand, if the estimation of the coefficients in $\hat{y}_1$ is good enough, then $\hat{y}_1 \approx y_1$, the unknown part of the spectrum $\hat{y}_2$ can be initialized to zero $\hat{y}_2^{(0)} = 0$,

and the CS procedure will be able to find the global minima. In this manner, the recursions are defined as,

$$\begin{cases} \hat{y}_2^{(0)} = 0, & k = 0, \\ \hat{y}_2^{(k)} = (1-\gamma_k)\,\hat{y}_2^{(k-1)} + \gamma_k\,(1-S)\left[T\left(\Phi\left(T^{-1}\left(\hat{y}_1 + \hat{y}_2^{(k-1)}\right)\right)\right) + \eta_k\right], & k \geq 1. \end{cases} \tag{4.7}$$

The different solutions studied here rely on two different approaches. First, the use of interpolation in Fourier space is proposed to obtain a better estimate of y_1. In this manner, the unknown Fourier coefficients contained in the gaps left by the discrete angular sampling are estimated, making the estimation of y_1 easier. Second, an initialization for the CS algorithm is suggested, providing a good starting point of y_2 for the CS reconstruction. Following, different methods are proposed to be used as an initialization of the CS procedure.

4.6.1 Initialization with Filtered Back-Projection

A straight solution is to obtain an initial Filtered Back-Projection (FBP) reconstruction and rely on the sampling operator S to segment the coefficients corresponding to the known projections correctly from the spectrum of the reconstructed image. However, this method is expected to fail, as the segmentation is not good enough since the location of the Fourier coefficients is not known accurately. In addition, FBP reconstruction does not try to compensate the missing wedge, and thus, initialization for the unknown coefficients should be zero, $\hat{y}_2^{(0)} = 0$. The flowchart of the procedure is shown in Figure 4.15.

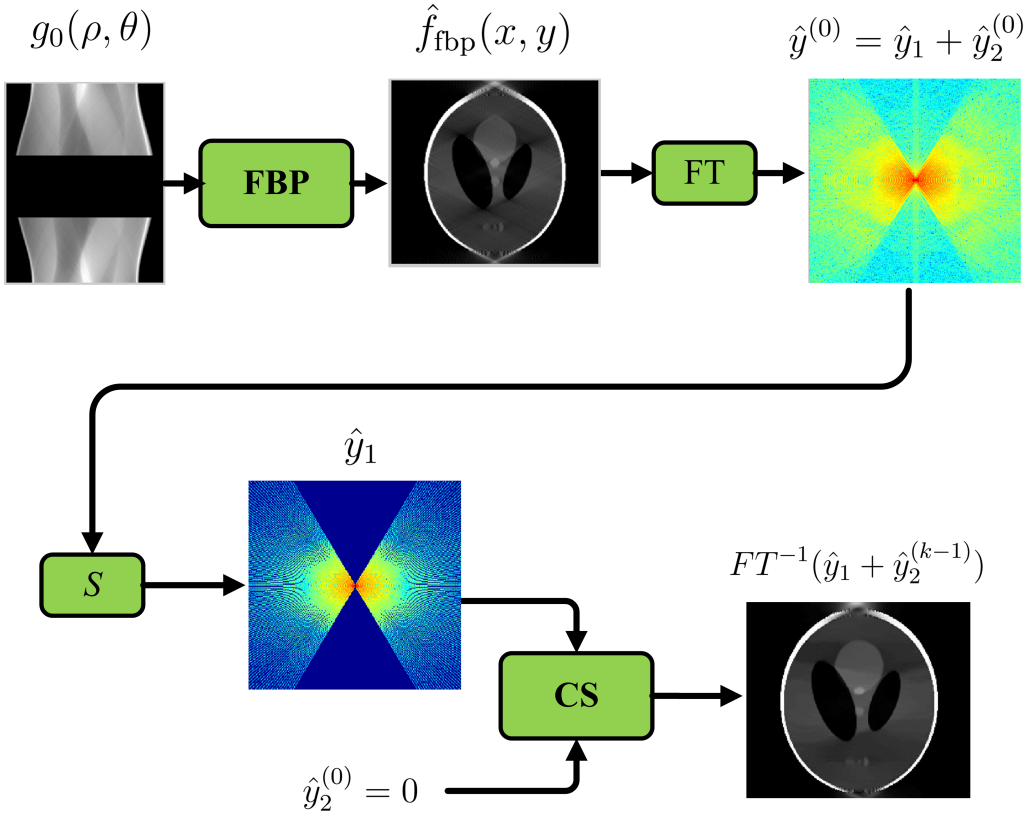


Figure 4.15: *Flowchart of the CS method with initialization by FBP.*

4.6.2 Initialization with Maximum Likelihood Expectation Maximization

A Maximum Likelihood Expectation Maximization (MLEM) reconstruction can be performed to define the intial image. The estimated known part, $\hat{y}_1$, is then segmented by S in the spectrum of the MLEM estimate. Then, instead of initializing the unknown portion of the spectrum with zero, a modification is here proposed by estimating these coefficients with the MLEM estimates. By MLEM, the most

likely image that caused the projections is obtained, and the missing coefficients corresponding to the missing wedge are estimated, providing a good initialization for $\hat{y}_2$, $\hat{y}_2^{(0)} = \hat{y}_2^{MLEM}$. In other words, the MLEM estimate serves as an initialization procedure for the CS method, see Figure 4.16.

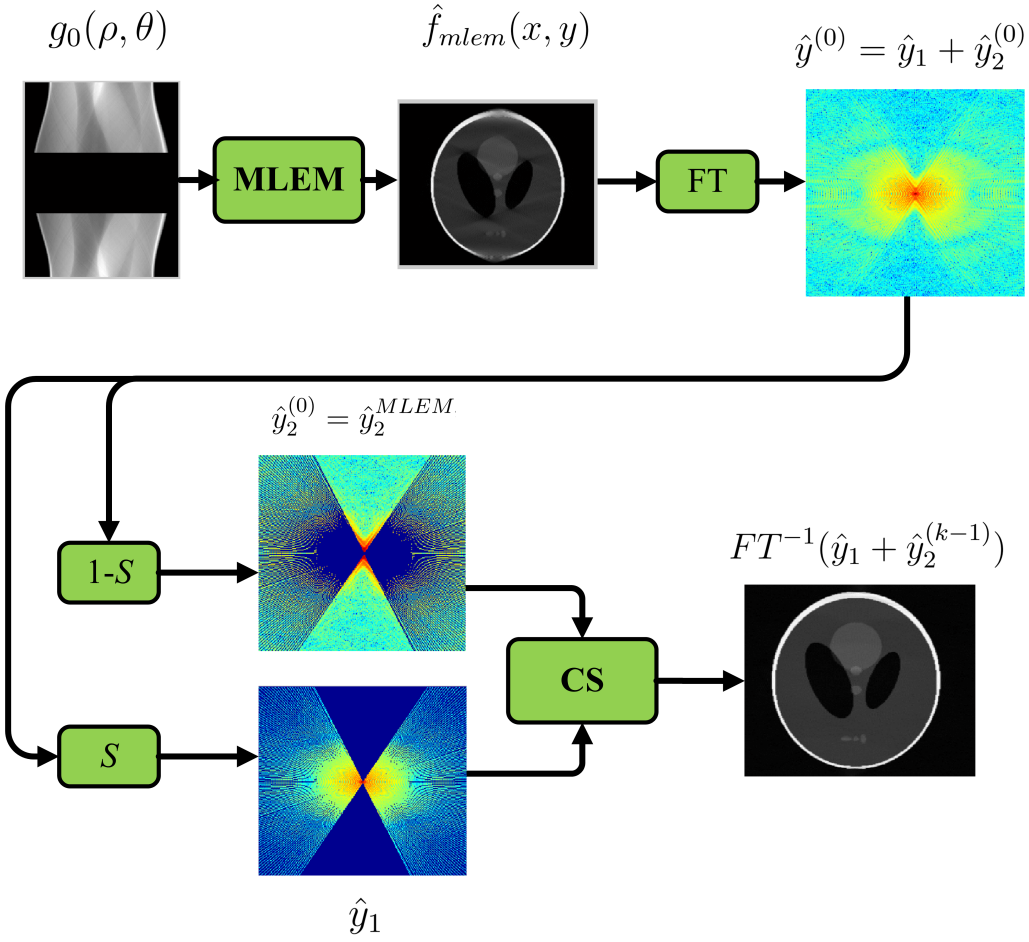


Figure 4.16: *Flowchart of the CS method with initialization by MLEM.*

4.6.3 Initialization with Interpolation in Fourier Space

The problem of exact segmentation of the FT coefficients is solved by using interpolation in Fourier space. This is performed by using Direct Fourier reconstruction (DFR) method as described in [36]. The missing Fourier coefficients in the gaps between the discrete projections (within the known range) are interpolated using the known projections and the Fourier-slice theorem. In this manner, $\hat{y}_1$ (the Fourier coefficients segmented by S) is a better approximation of the original y_1. By using DFR and simple inverse Fourier Transform of the assembled Fourier space, an initialization of the image and its spectrum is obtained, see Figure 4.17. In this initial procedure, the method leaves the missing wedge, and thus $\hat{y}_2^{(0)} = 0$, which should

be later filled by the CS approach.

Sinogram Gap-Filling with Inpainting

In addition, if some initial estimation for the missing projections lying within the missing wedge is found, these projections will contribute in the oversampling in the low frequencies, contributing to a better estimate of $\hat{y}_1$. Thus, instead of leaving the missing wedge empty in the initialization, inpainting can be used to fill the gap in the sinogram. The inpainted region provides a low frequency approximation of the missing wedge. Then, DFR can be used to assemble an estimated Fourier spectrum. In this manner, interpolation provides a better estimate for $\hat{y}_1$, and the unknown coefficients in Fourier space caused by the missing wedge are initialized by those estimated by the inpainting method, $\hat{y}_2^{(0)} = \hat{y}_2^{inpainted,DFR}$, see Figure 4.17. However, note that if the estimation of the coefficients corresponding to the known part $\hat{y}_1$ is good enough, the unknown part could be initialized to zero $\hat{y}_2^{(0)} = 0$, and the CS procedure still should obtain a good estimate of the image.

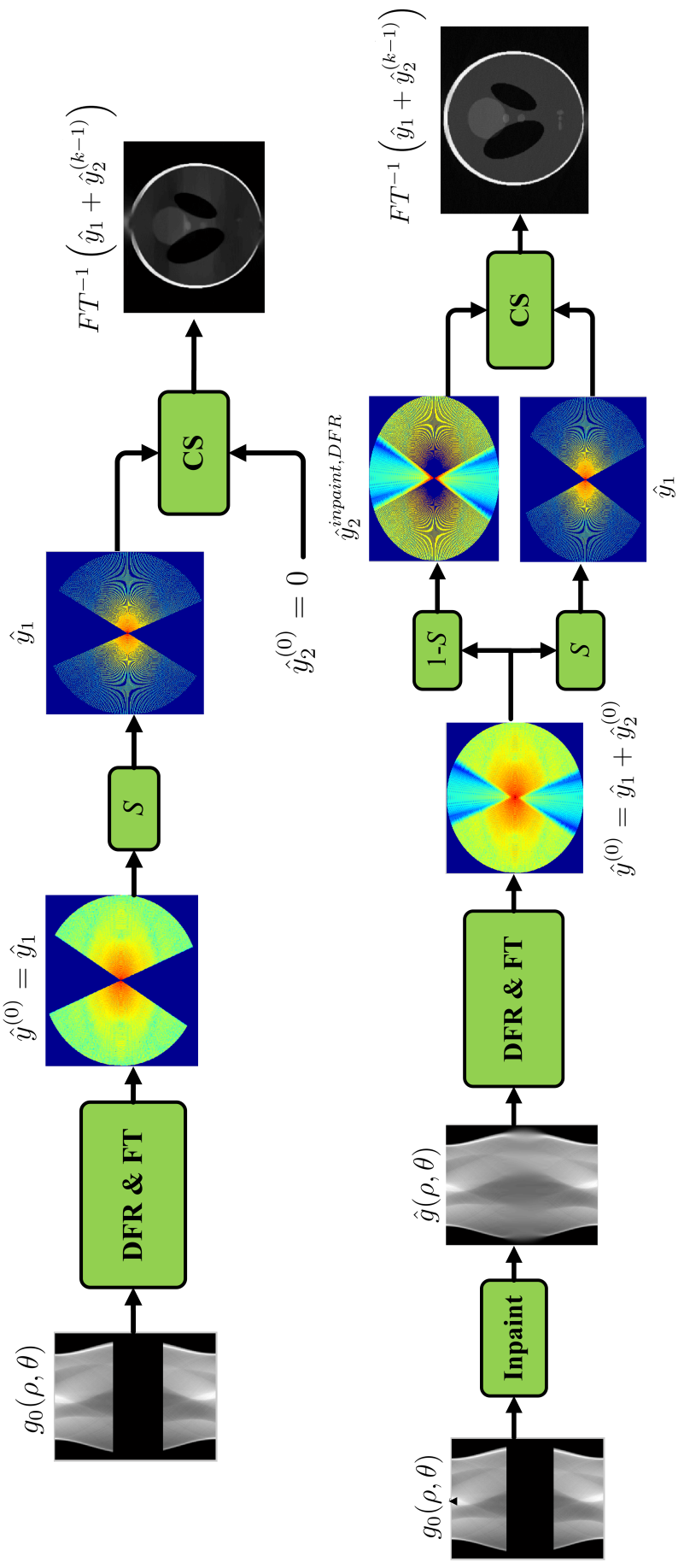


Figure 4.17: *Up: Flowchart of the CS method with interpolation in Fourier space. Down: Flowchart of the CS method with sinogram gap-filling by inpainting and DFR.*

Sinogram Gap-Filling Using Maximum Likelihood Expectation Maximization

As mentioned, MLEM obtains an estimate of the most likely image that caused the measured projections. Thus, this image can be used to obtain an initial estimation of the missing projections, that is, the projections corresponding to the missing wedge. These estimated projections can be used to fill the gap in the original sinogram. In this manner, the original sinogram $g_0(\rho, \theta)$ is used to obtain an estimate by MLEM of the original image $\hat{f}(x, y)$. Then, the projections of $\hat{f}(x, y)$ in the angles corresponding to the missing wedge are obtained, and thus an initial estimate for the gap $\hat{g}_{gap}(\rho, \theta)$ is created. The estimated missing wedge is then used to fill the gap in the original sinogram. A complete estimated sinogram $\hat{g}(\rho, \theta)$ is obtained. By using $\hat{g}(\rho, \theta)$, it is possible to use interpolation in Fourier space to estimate the complete spectrum by means of DFR, see Figure 4.18.

The interpolation in Fourier space in the DFR method using the estimated complete spectrum produces a better estimate for $\hat{y}_1$. The unknown portion, $\hat{y}_2^{(0)}$ is initialized by the estimated spectrum $\hat{y}_2^{(0)} = \hat{y}_2^{MLEM,DFR}$ or, assuming that $\hat{y}_1 \approx y_1$, it can be initialized to zero, $\hat{y}_2^{(0)} = 0$.

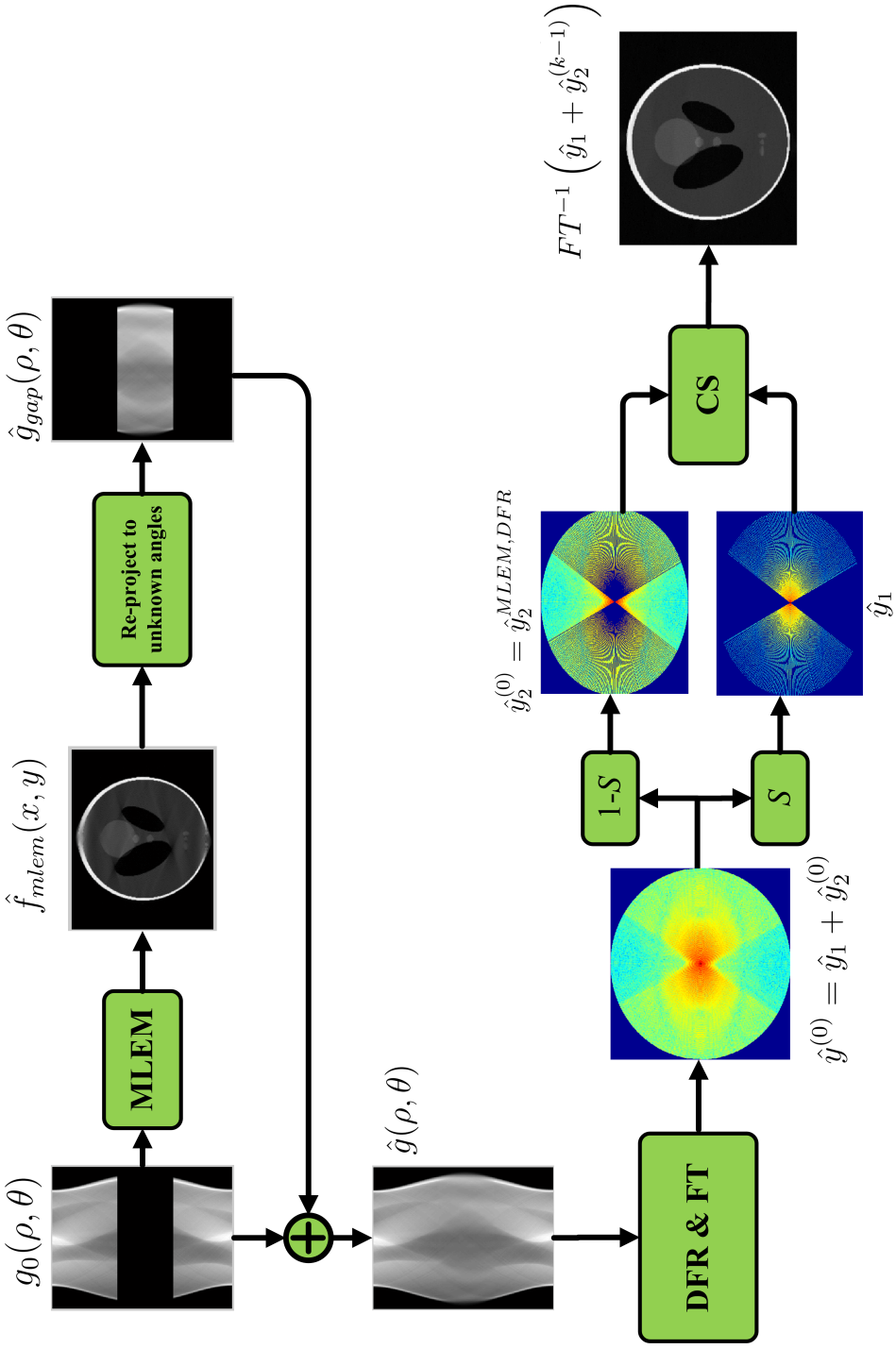


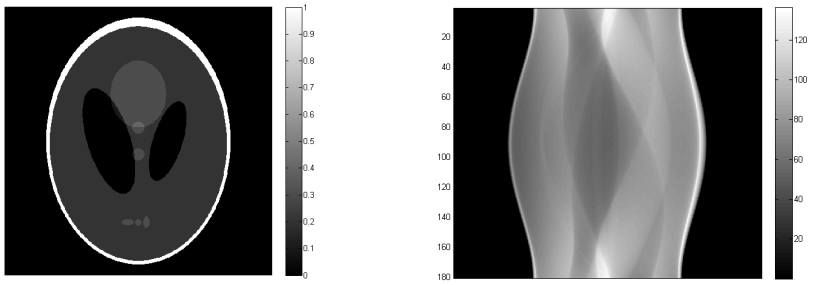
Figure 4.18: *Flowchart of the CS method with initialization by sinogram gap-filling using MLEM and DFR.*

5. EXPERIMENTAL DATASETS

In order to assess the proposed methods, numerical phantom data are used. In this manner, with the numerical phantom it is possible to perform quantitative evaluation of different reconstruction methods or algorithms. In this chapter, two different datasets are introduced. First the numerical phantom, which is used as a ground truth is introduced: the Shepp-Logan head phantom. Following, a brief description of the real TEM data used to perform tomographic reconstruction is presented.

5.1 Shepp-Logan Phantom

The Shepp-Logan phantom is a well-known numerical model resembling a human head section designed by L. A. Shepp and B. F. Logan in 1974 [67]. It consists of ellipses representing the attenuation the x-rays suffer across the different brain structures, see Figure 5.1.



(a) *(b)*

Figure 5.1: *The Shepp-Logan head phantom and its sinogram.*

This phantom is not the best synthetic model to use when simulating ET since it is a model of the attenuation of x-rays across macroscopic object. On the other hand, ET deals with microscopic objects in nanometric scale, and the electron-matter interaction is based on different physical principles than in, for instance, CT or PET imaging. Even though physical principles are different, the mathematics of reconstruction rely on projections and, as has been discussed in previous chapters, TEM images can be considered as projections of some function of the specimen's structure.

Moreover, the Shepp-Logan head phantom is a benchmark in medical image reconstruction from projections, and it is widely used in scientific publications to test and compare reconstruction methods, including limited angle tomography. In addition, further motivations arise from the lack of an established numerical phantom used as a benchmark in ET. Some phantoms can be found in literature, although they are focused on some concrete applications such as the mitochondria phantom designed in [43]. Considerable amount of algorithms and methods focused on limited angle tomography that are tested by using the Shepp-Logan phantom also suggests the use of it. Finally, this artificial dataset is a 2D object, and the initial simulations performed were 2D reconstructions. However, 2D reconstructions can be easily extended to 3D.

5.1.1 Alternative Phantom

It was mentioned that the missing wedge causes the loss of structures perpendicular to the plane defined by the electron beam and the tilt axis [18]. In order to illustrate this effect, an alternative phantom image obtained from [28] is used. This image, shown in Figure 5.2, contains planar structures that are lost with the missing wedge.

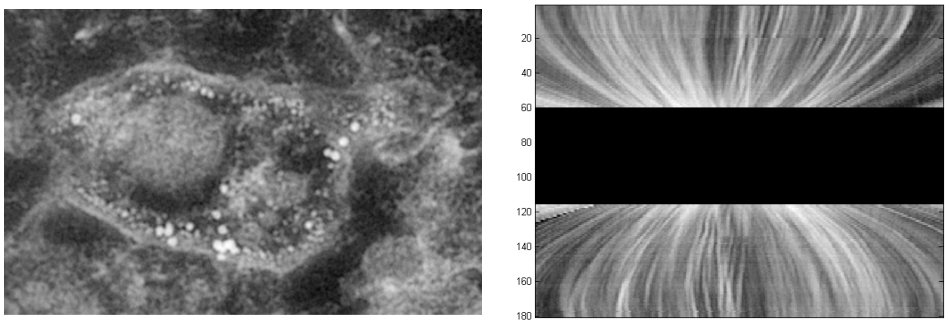
(a) Sample image used in [28]. *(b) Modification*

Figure 5.2: *Alternative phantoms used to illustrate the missing wedge effect.*

A modification of this phantom is also shown in Figure 5.2. It includes, instead of planar structures, a variation with curved structures, intuitively more similar to real structures present in biological specimens.

5.2 Transmission Electron Microscope Dataset

In addition to the simulations with the artificial data, the approaches were tested using real TEM data. The dataset used, see Figure 5.3, corresponds to a series of TEM images obtained with the JEM-2100F Field Emission Electron Microscope, at a voltage of 200 kV. The dataset contains a series of 124 images of a cellular vesicle, taken at an interval of approximately one degree in the range [-65°, +58°]. The magnification of images is 10000, with a size of 712x958 pixels and a resolution of 1×1 nm/pixel.



(a) *(b)*

Figure 5.3: *Real dataset used for reconstruction. (a) Zero tilt TEM image from the data set. (b) Section of the sinogram with the gap caused by the missing wedge.*

5.3 Evaluation Methods

The obtained results with the different reconstructions need to be evaluated. For this reason a measure of the success of the reconstructions is needed. Visual evaluation can be a good starting point to recognize how a reconstruction method performs, and for this reason it was used for preliminary results. However, in order to be more precise, quantitative measures are needed. A numerical phantom (the Shepp-Logan phantom) was used as a ground truth in the tests. An assessment method to measure the improvement caused by the gap-filling method is to use the mean squared error (MSE) normalized by the MSE obtained in a reconstruction where no gap-filling was used. Thus, the %MSE is used as an assessment criteria [62]

$$\%\mathrm{MSE}\left(f_{\mathrm{est}}\left(x,y\right),f_{\mathrm{gap}}\left(x,y\right)\right)=\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left[f_{\mathrm{est}}\left(x,y\right)-f_{\mathrm{orig}}\left(x,y\right)\right]^{2}}{\sum_{x=1}^{M}\sum_{y=1}^{N}\left[f_{\mathrm{gap}}\left(x,y\right)-f_{\mathrm{orig}}\left(x,y\right)\right]^{2}}\times 100, \quad (5.1)$$

where the f_{est} denotes the reconstruction with gap-filling (missing wedge correction), f_{gap} is the reconstruction where the missing wedge is left empty and f_{orig} is the original numerical phantom. The normalizing image f_{gap} is a standard Filtered

Back-Projection (FBP) reconstruction obtained from the incomplete sinogram. In this manner, if the %MSE for a given reconstruction is smaller than 100, it means that the result obtained is better than FBP reconstruction. It gives a measure of improvement of the gap-filling approach used with respect to FBP. This means that the lower the %MSE value is the better is the reconstruction, where a value of 0 for the %MSE means perfect reconstruction.

In practical reconstruction with real data, a different version of Mean Squared Error (MSE) is used as a heuristic rule to stop iterations. The MSE between consecutive iterations is calculated

$$\mathrm{MSE}_{(k)}^{(k+1)} = \mathrm{MSE}\left(f_{\mathrm{est}}^{k}, f_{\mathrm{est}}^{k+1}\right), \tag{5.2}$$

and when the difference between two consecutive measures drops below a determined level c, $\mathrm{MSE}_{(k)}^{(k+1)} - \mathrm{MSE}_{(k+1)}^{(k+2)} < c$, the iterations stop.

6. RESULTS

Several methods were proposed and described in this work. The simulations of the mentioned methods were carried out by using the artificial data described in the previous chapter. In this chapter, the results obtained in the simulation different methods are presented. Next, the reconstruction using the real data set is given.

Initial simulations were performed by using the Shepp-Logan phantom. The first logical step is to reconstruct and inspect the results by using reconstruction methods without gap-filling. In Figure 6.1 the results obtained from FBP reconstruction are displayed, both by using a complete sinogram and an incomplete sinogram simulating the missing wedge.

(a) *(b)* *(c)* *(d)* *(e)*

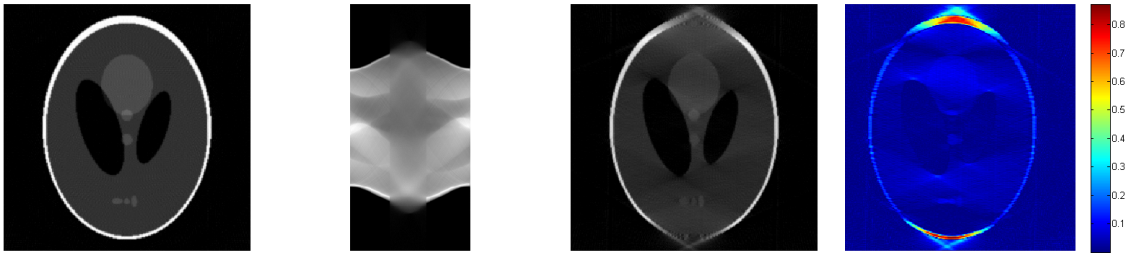
Figure 6.1: *FBP reconstruction of the head phantom. (a) Original complete sinogram. (b) FBP from complete sinogram. (c) Incomplete sinogram due to the missing wedge. (d) FBP reconstruction of (c). (e) Difference between (b) and (d).*

The angular coverage is the same as in the real data set, consisting of projections in the range of [-65°, +58°]. Reordering and representing it in the range $[0°, +180°)$ the angular coverage is equivalent to $[0°, +58°] \cup [+115°, 180°)$, leaving a missing wedge of 57° in the middle of the sinogram, Figure 6.1. This is the usual case in electron tomography (ET), where the missing wedge is of approximately 60°. The missing wedge causes loss of resolution and elongation of the image features in the vertical direction of the image as can be seen in Figure 6.1.

It is needed to mention that the first attempts to reconstruct the simulated data set were non exhaustive. The main purpose was to experimentally inspect and understand how the missing wedge affects the reconstructions. Not only this, but also learn how the different reconstruction methods and proposed solutions affects the missing wedge and thus try to steer the reconstruction towards a correct solution.

6.1 Iterative Projection - Back-projection

The iterative procedure described in Figure 4.6 in subsection 4.3.1 was applied. It was already argued that the use of powerful constraints in ET is not possible due to the complexity of real biological samples. In these initial simulations, the used head phantom could suggest the use of some constraints such as piecewise smoothness, or limited extent. However, in order to be as close as possible to the real case, these constraints were not applied, leaving only minor corrections on the non-negativity and filtering in image domain using a Gaussian smoothing kernel in an attempt to alleviate artifacts in reconstruction.



(a) *(b)* *(c)* *(d)*

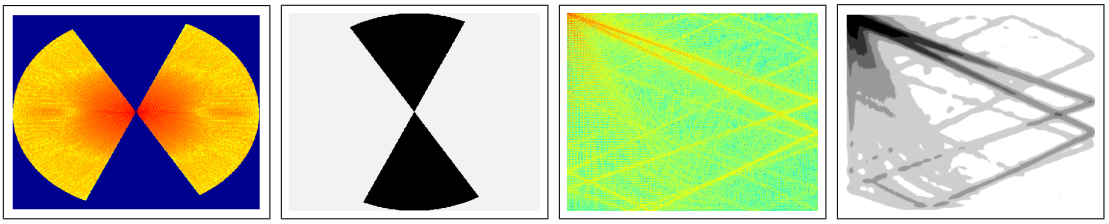
Figure 6.2: *Reconstruction of the head phantom applying iterative projection-back-projection method shown in the flowchart in Figure 4.6.(a) FBP from complete sinogram. (b)Filled sinogram. (c) Reconstructed image. (d) Difference between (a) and (c).*

The result of applying some iterations of the mentioned procedure are shown in Figure 6.2. It can be seen that the gap is filled somehow, providing some continuity to the sinogram, although in image domain no big improvements can be seen except some minor reduction in the maximum absolute error.

6.2 Gap-Filling Using a Dedicated DCT-Domain Filter

As it was described in subsection 4.3.2, this method consists of creating a DCT-domain filter to interpolate the gap in Fourier space caused by the missing wedge, by using the method shown in Figure 4.8. The first step is then to create the Fourier space based on the projections of the sinogram by using the direct Fourier reconstruction method as presented in the image reconstruction from projections chapter. The assembled Fourier space is shown in Figure 6.3, where the characteristic shape of the missing wedge is recognizable. The following steps, shown in Figure 6.3 is to extract the shape of the gap, and its characteristic DCT coefficients are used to create the dedicated DCT-filter.

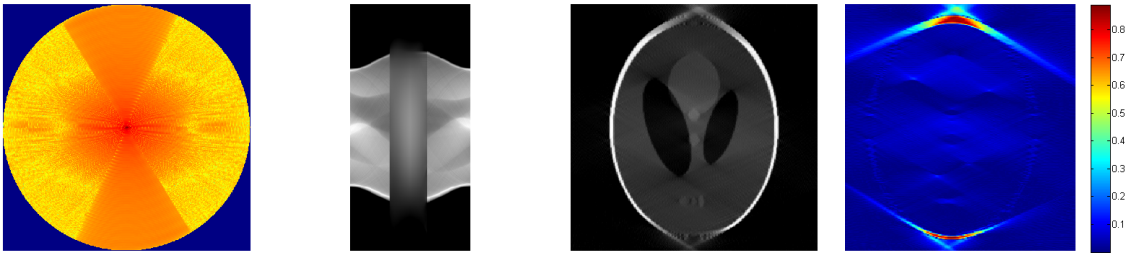
The generated dedicated filter is then iteratively applied and the results obtained for the simulated phantom are shown in Figure 6.4. Inspecting the results it can be



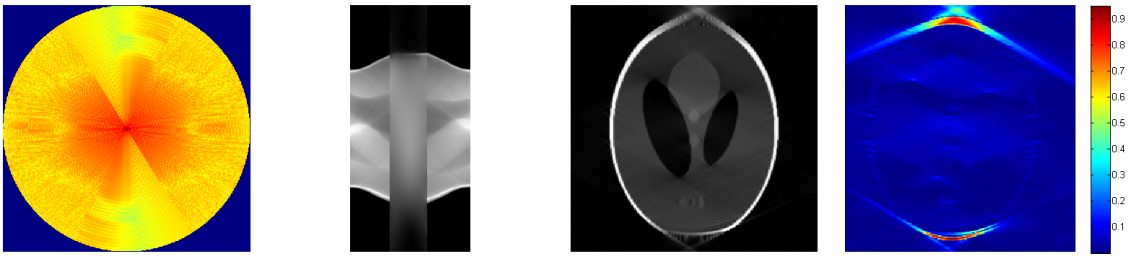
(a) *(b)* *(c)* *(d)*

Figure 6.3: *Process of extracting the DCT dedicated mask. (a) Fourier space with missing wedge. (b) Mask for the missing wedge. (c)DCT coefficients of (b). (d) Dedicated filter for the missing wedge using (c).*

seen that, as expected, the missing wedge in Fourier space was filled. However, these interpolated Fourier coefficients are not consistent with their neighboring known coefficients and although the gap was filled in Fourier space, in the sinogram it consist of irrelative information.



(a) Filled Fourier space *(b) Filled sinogram* *(c) Reconstruction* *(d) Differences*



(e) Filled Fourier space *(f) Filled sinogram* *(g) Reconstruction* *(h) Differences*

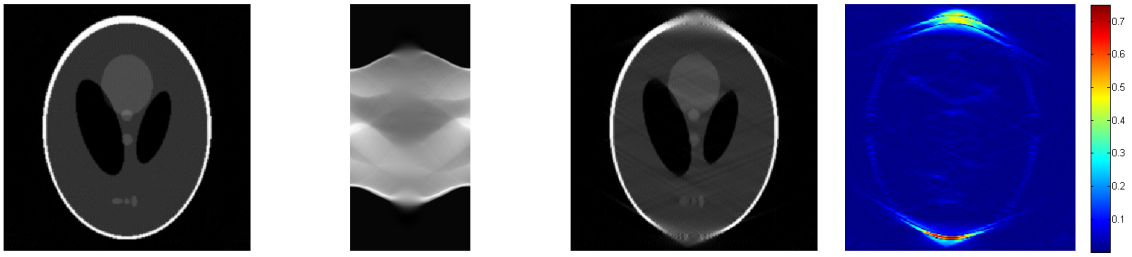
Figure 6.4: *Results of applying a dedicated DCT filter in Fourier space using method shown in Figure 4.8. (a)-(d) After 50 iterations. (e)-(h) After 300 iterations.*

This method does not seem to work well when applied in Fourier space. In order to analyze these undesired results, one has to go to the motivation for the use of DCT. As explained, DCT has higher energy compaction for highly correlated data. Thus, it is expected to work well the spatial domain, as normally neighboring pixels will be correlated. However in Fourier space data is decorrelated, so consequently Fourier coefficients will have low correlation. That is why using this kind of method

in Fourier space is not wise, and thus failure is expected.

6.3 Image Inpainting

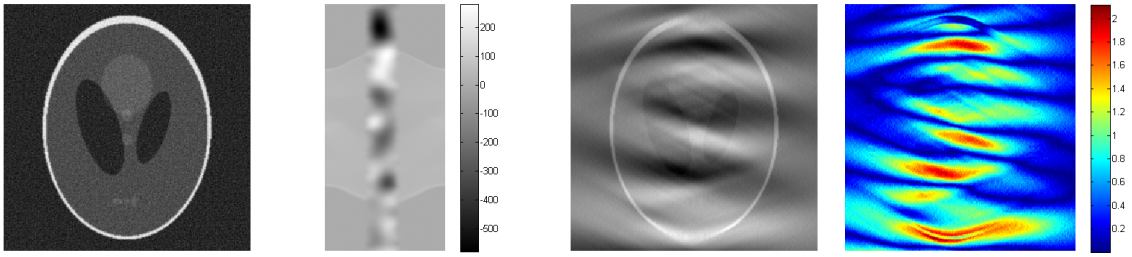
This method gives quite interesting results as it treats the sinogram as an image and then tries to interpolate the missing values as professional painters would do in image restoration. The application of this method in interpolating the missing parts in the sinogram is not new and it was proven to give acceptable results in [62], although it was used to interpolate small gaps in the sinogram. The results of applying inpainting to fill the missing wedge in the sinogram are shown in Figure 6.5. These results were promising at the first glance as it seemed to provide continuity in the sinogram and smooth the artifacts of the gap in the image domain.



(a) *FBP reconstruction from complete sinogram.* ***(b)*** *Inpainted sinogram.* ***(c)*** *Reconstruction using (b).* ***(d)*** *Differences between (a) and (c).*

Figure 6.5: *Inpainting in the sinogram.*

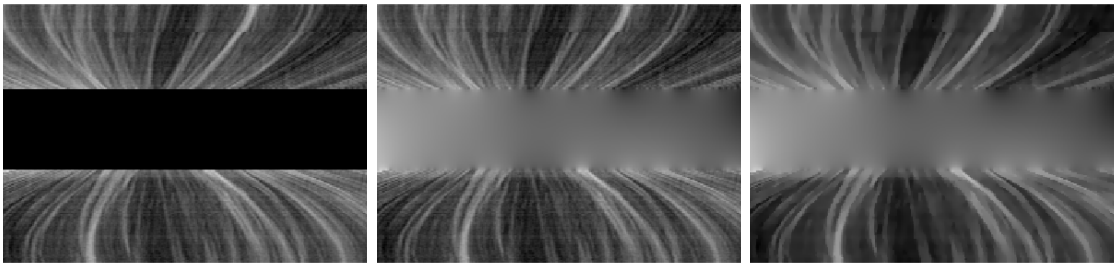
However, further simulations closer to real case resulted in failed reconstructions. As shown in Figure 6.6, if the sinogram is corrupted with additive Gaussian noise the inpainting method fails. The inpainted region in the sinogram diverges providing totally inaccurate results as can be seen in the reconstructed images.



(a) *FBP reconstruction from complete noisy sinogram.* ***(b)*** *Inpainted sinogram.* ***(c)*** *Reconstruction using (b).* ***(d)*** *Differences between (a) and (c).*

Figure 6.6: *Inpainting applied to a noisy sinogram.*

If the inpainting method is applied to real data, the obtained results are not as expected. This is shown in Figure 6.7, where the inpainting was performed by using the sinogram of the real dataset. In an attempt to overcome possible problems due to the noise, additional filtering was used on the sinogram. Concretely, anisotropic diffusion filtering was applied to the original sinogram. In this manner, the sinusoidal trajectories of the image features were well preserved in the sinogram and thus the inpainting should follow these paths within the gap. Still, no improvements seem to be achieved, as can be seen in Figure 6.7(c).



(a) *Original sinogram.* ***(b)*** *Inpainted sinogram* ***(c)*** *Inpainting of the filtered sinogram.*

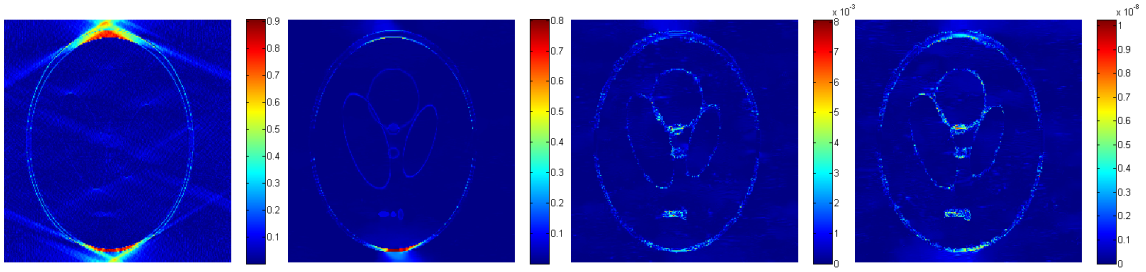
Figure 6.7: *Inpainting performed on real data. The sinogram in (c) was filtered using anisotropic resolution.*

It is necessary to note that the gap caused by the missing wedge is too large to be correctly filled by the image inpainting in real cases. In cases where the gaps are smaller and completely surrounded by data, inpainting could successfully interpolate the missing values. This approach does not seem to be a correct one, although if better inpainting algorithms are designed they could be applied yielding better results.

6.4 Compressed Sensing Approach

The last set of simulations were performed with the CS approach defined in [64] and with the modifications of this procedure which were proposed previously. A demonstration software implemented by Egiazarian *et al.* can be found in [68], and it was used for some of the initial simulations. The proposed variations of the method were implemented by modifying the demonstration software.

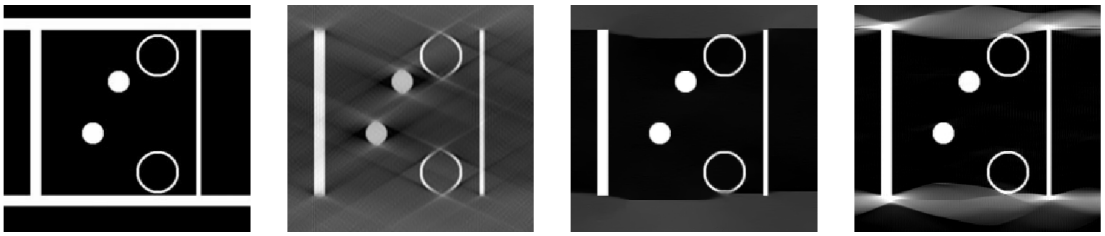
(a) *Initial reconstruction* ***(b)*** *1000th iteration* ***(c)*** *10000th iteration* ***(d)*** *50000th iteration*



(e) ***(f)*** ***(g)*** ***(h)***

Figure 6.8: *Simulations performed with the original method in [64], and described in Figure 4.12, for 0, 1000, 10000 and 50000 iterations in (a)-(d). In (e)-(h) the reconstructions error with respect the original phantom are shown. Note the error drops several orders of magnitude for high number of iterations, in (g) and (h), showing reconstruction close to perfection.*

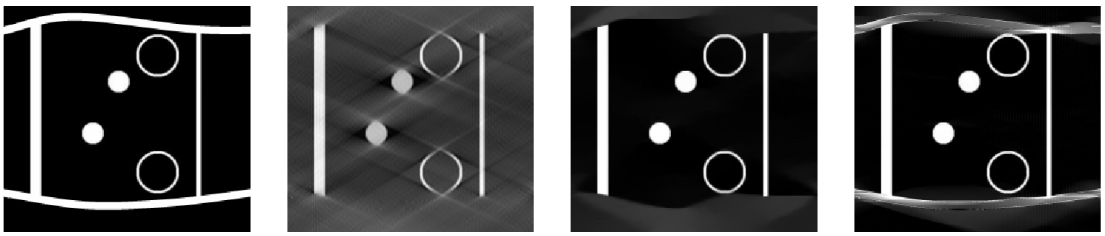
Initial simulations as described in the original procedure [64] yielded exact reconstruction for the head phantom, as can be seen in Figure 6.8. However, in some cases the algorithm was not able to recover completely the original object. This is the case if structures are exclusively contained in the direction of the missing wedge. Even though the CS reconstruction compensates the artifacts caused by the missing wedge, the structures parallel to the direction of the missing wedge are not recovered. This effect can be seen in the reconstruction of the alternative phantom shown in Figure 6.9. This reconstruction, is then the best possible estimate given by the CS approach for this type of images.



(a) Original image *(b) FBP with missing wedge* *(c) CS algorithm after 1000 iterations* *(d) CS algorithm after 30000 iterations*

Figure 6.9: *Reconstruction of the alternative phantom from [28] using the CS simulations of [64]. The missing wedge is in the x-axis direction. Structures parallel to this direction are completely lost in the reconstruction.*

However, line-type structures are not usually present in real biological samples, where the membranes are of irregular shape, and thus they can be partially recovered. The reconstruction of the modified alternative phantom, closer to a real case, is shown in Figure 6.10. The line-type structures affected were substituted by curved ones, representing membranes contained within the direction of the missing wedge. In this case, even though reconstruction is not perfect, the CS reconstruction is able to partially resolve these membranes.



(a) Original image *(b) FBP with missing wedge* *(c) CS algorithm after 1000 iterations* *(d) CS algorithm after 30000 iterations*

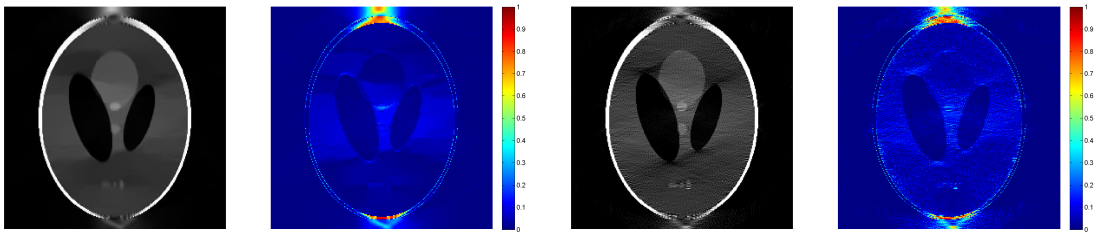
Figure 6.10: *Reconstruction of the modified alternative phantom from [28] using CS simulations of [64]. The missing wedge is in the x-axis direction. Now, structures are not completely unresolved in the missing wedge direction.*

6.5 Modifications of the Compressed Sensing Approach

Previous reconstructions are simulations as performed in [64]. However, note that this approach simulates the Radon projections by taking radial lines in the complete Fourier space of the original image. That is, the reconstruction is based in a sparse set of known Fourier coefficients whereas, in a more realistic situation, this is not possible, as the acquired data corresponds directly to real imperfect projections of the original object. In other words, the available data is the incomplete sinogram. This is why a modification of this reconstruction method was needed.

6.5.1 Initialization by Filtered Back-Projection

A straight-forward approach would be direct application of the CS algorithm using an FBP reconstruction. The FBP reconstructed image was used to obtain $\hat{y}_1$ with S and the unknown coefficients are set to zero, $\hat{y}_2^{(0)} = 0$. However, this method does not completely work, as it was previously discussed. The results are shown in Figure 6.11, where the best result achieved in the iterative reconstruction (the lowest %MSE) and its reconstruction error are shown. In addition, the reconstruction after 25000 iterations of the CS procedure is also shown, giving an idea of the convergence of the algorithm. Inspecting the Figure 6.11, it can be seen that that the reconstruction still presents artifacts due to the missing wedge.

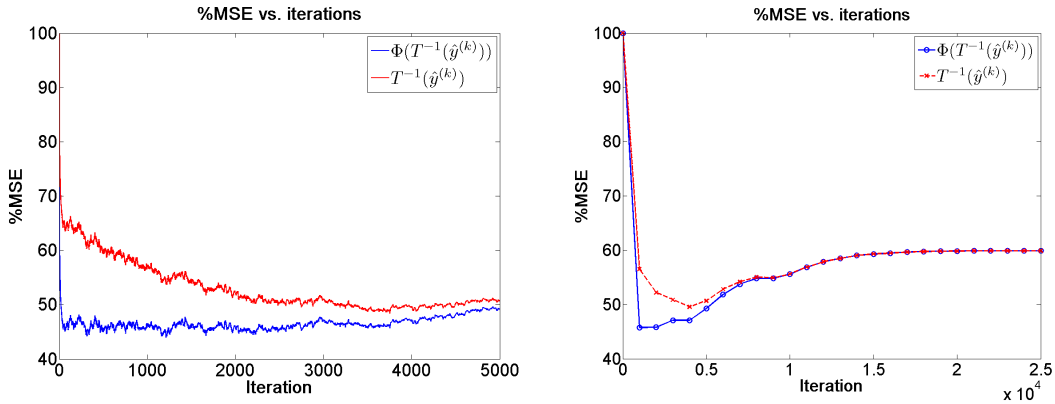


(a) *1211th iteration* ***(b)*** *reconstruction error of (a)* ***(c)*** *25000th iteration* ***(d)*** *reconstruction error of (c)*

Figure 6.11: *Results of applying CS algorithm directly over a FBP reconstruction of the head phantom with missing wedge. The method improves the normal FBP reconsutruction, although it does not compensate completely the missing wedge.*

In Figure 6.12, the %MSE progression curves in the iterations are shown to visualize the reconstruction process. The %MSE figures contain two curves. One curve corresponds to the %MSE of the image estimate in the k-th iteration $T^{-1}\left(\hat{y}^{(k)}\right)$, which is obtained after inverse Fourier transformation of the estimated spectrum. The second curve corresponds to the denoised version of the estimate $\Phi\left(T^{-1}\left(\hat{y}^{(k)}\right)\right)$, which will be preferred as reconstruction estimate since it lacks of the excitation noise, and consequently presents a better %MSE. Note, however, that both curves converge to the same value as iterations progress. Different versions of the sampling operator S were used in the simulations. However, the simulation results shown in the following subsections were obtained using a filtered version of the original mask S. The effects of S on the reconstruction are discussed later.

Inspecting the results shown in Figure 6.12, it can be seen that the %MSE values are always under 100%, meaning that the method improves the reconstruction with respect to the FBP reconstruction without missing wedge correction. The reconstructions start with good results, with the best %MSE achieved within the first 2000 iterations and a minimum present in 1211 iterations. However, as iterations



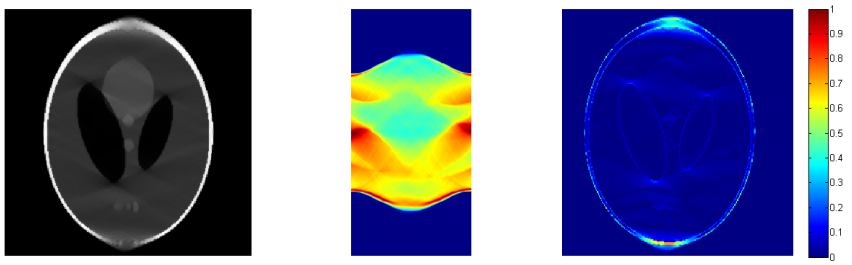
(a) First 5000 iterations *(b) First 25000 iterations*

Figure 6.12: *%MSE results for CS algorithm applied by using an FBP reconstruction. The results show improvement with respect to the FBP reconstruction.*

progress, it can be seen that the reconstruction becomes degraded and the %MSE converges to a value of approximately 60%, result that it is already better than the initial FBP reconstruction.

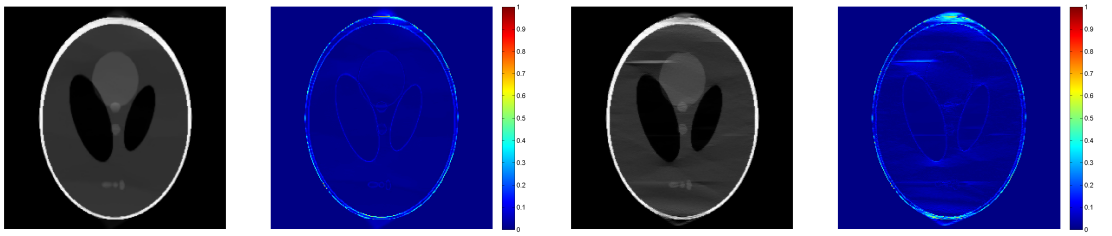
6.5.2 Initialization by Maximum Likelihood Expectation Maximization

As mentioned in chapter 5, an approach to overcome the problems present in direct application of CS method is to use maximum likelihood expectation maximization (MLEM) to estimate the unknown portion of the sinogram. An initial estimate of the image is obtained and then used to initialize the CS approach. The algorithm used is Ordered Subsets Expectation Maximization (OSEM), and the obtained results are shown in Figure 6.13. The reconstruction obtained by using OSEM is already an improvement compared to the normal FBP reconstructions, although still some of the artifacts caused by the missing wedge are present.



(a) *(b)* *(c)*

Figure 6.13: *Reconstruction using MLEM with ordered subsets.*



(a) 7000th iteration *(b) reconstruction error of (a)* *(c) 25000th iteration* *(d) reconstruction error of (c)*

Figure 6.14: *Results of applying CS algorithm using OSEM initialization.*

The CS reconstruction is applied with the initialization $\hat{y}_2^{(0)} = \hat{y}_2^{MLEM}$. The results are shown in Figure 6.14. The best estimate is obtained with 7000 iterations, and it has improved the image estimate obtained by OSEM reconstruction in Figure 6.13. Again, it can be seen that the %MSE achieves a minimum around 7000 iterations and then it increases converging to a worse result, see Figure 6.15. Although the missing wedge effect was corrected up to a good level, the reconstruction presents deterioration as iterations progress due to some new artifacts that appear in the image, as seen in Figure 6.14, possibly due to a bad estimate of $\hat{y}_1$.

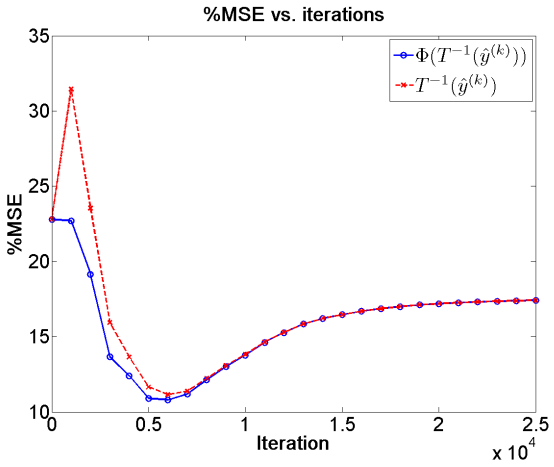
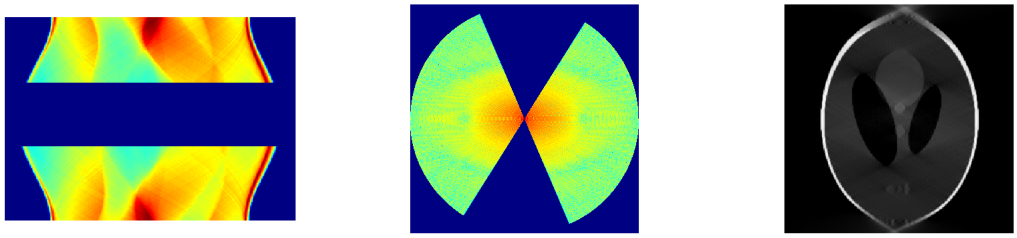


Figure 6.15: *%MSE results for CS algorithm CS algorithm using OSEM initialization.*

6.5.3 Initialization Using Interpolation in Fourier Space

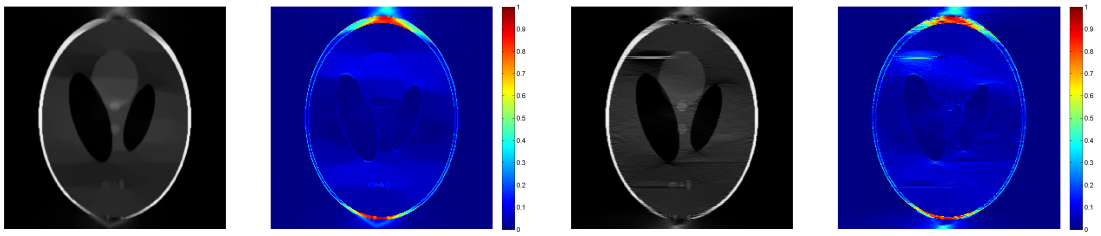
A proposed solution improve the results obtained with the CS method using the FBP reconstruction is to use Direct Fourier Reconstruction (DFR). In this manner, the available projections are used to interpolate the missing Fourier coefficients to obtain $\hat{y}_1$ and then the CS algorithm estimates the image. The interpolated Fourier space within the known range of projections and the obtained image are shown in Figure 6.16.



(a) incomplete sinogram *(b) Assembled Fourier space* *(c) Reconstructed image*

Figure 6.16: *DFR method with the missing wedge. The incomplete sinogram in (a) is used to assemble the Fourier Space (b) leaving the missing wedge empty. (c) Obtained image.*

After the DFR method is employed, the CS approach is used to improve the results, as shown in the flowchart in Figure 4.17. The results obtained with this method are shown in Figure 6.17. If the %MSE values in Figure 6.18 are analyzed, after the first 1000 initial iterations, the results are better than the reference FBP and the initial direct Fourier Reconstruction method. However, this approach does not improve the results obtained by applying CS method after FBP reconstruction.



(a) 1000th iteration *(b) reconstruction error of (a)* *(c) 25000th iteration* *(d) reconstruction error of (c)*

Figure 6.17: *Results of applying CS algorithm directly over a DFR reconstruction of the head phantom with missing wedge.*

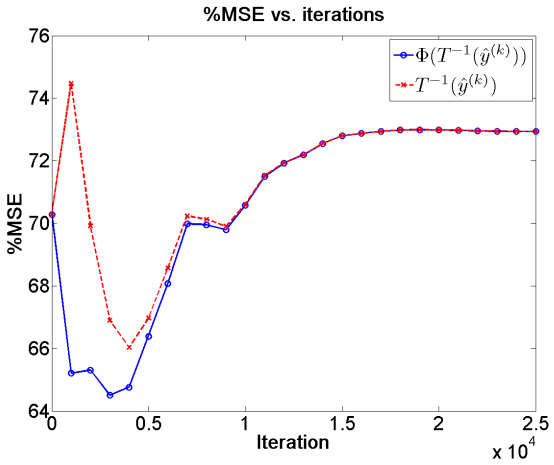
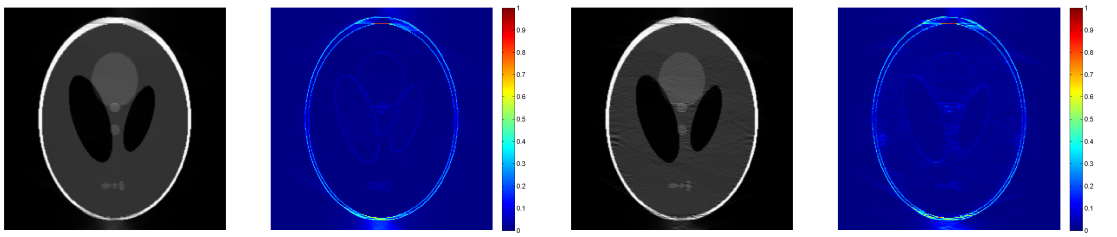


Figure 6.18: *%MSE results for CS algorithm applied using a DFR reconstruction. The results show some improvement with respect to the FBP reconstruction.*

Sinogram Gap-Filling Using Inpainting

Image inpainting can be used to estimate the missing wedge. Inpainting results gave interesting results provided the sinogram is noise-free, but failed in the case of the real data and noisy sinograms. Nevertheless, for the purpose of testing and understanding the proposed procedures better, inpainting is used for these simulations. By inpainting the gap in the sinogram, it is possible to obtain an initial guess of the missing wedge. Then, the initial estimate of the image (and its spectrum) is obtained by using DFR, method that interpolates the missing coefficients between the lines in Fourier space. After that, the CS method is then applied to complete the reconstruction, as shown in Figure 4.17.



(a) *7000th iteration* ***(b)*** *reconstruction error of (a)* ***(c)*** *25000th iteration* ***(d)*** *reconstruction error of (c)*

Figure 6.19: *Results of applying CS algorithm directly over a DFR reconstruction with initialization by inpainting.*

The results are shown in Figure 6.19. The best result were obtained by initializing $\hat{y}_2$ to zero, $\hat{y}_2^{(0)} = 0$, and are surprisingly good. The reconstruction achieves an image estimate where the missing wedge artifacts were reduced considerably, as can be seen in Figure 6.19. The best estimate is obtained with approximately 7000 iterations,

see Figure 6.20, with a %MSE of approximately 16.5%, although the reconstruction converges to a poor estimate.

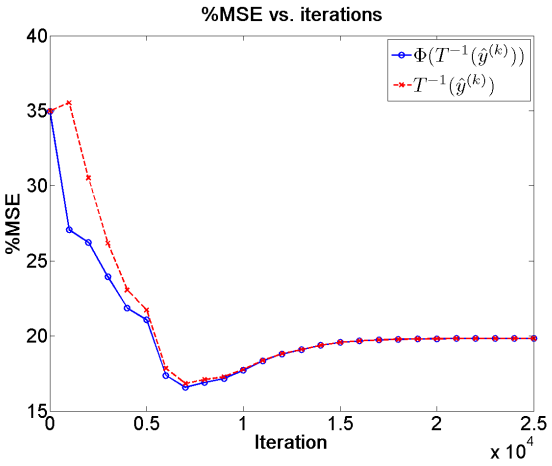
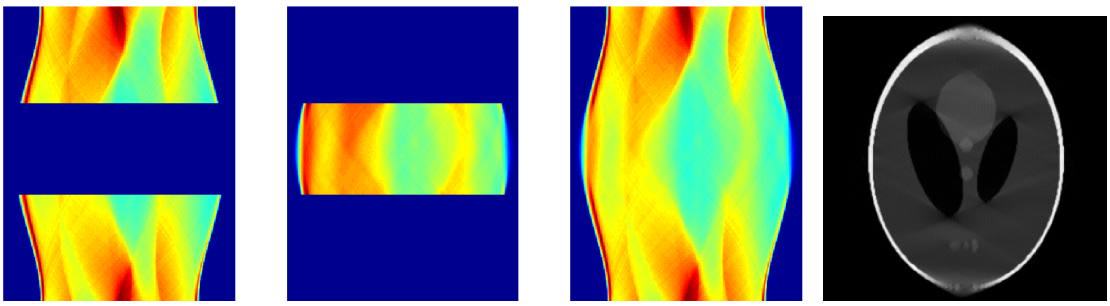


Figure 6.20: *%MSE results for CS algorithm using an inpainted DFR reconstruction.*

Sinogram Gap-Filling Using Maximum Likelihood Expectation Maximization

The last proposed method consists of estimating the gap in the original sinogram by the MLEM reconstruction. Using the OSEM approach, estimates of the missing projections were calculated and inserted into the original sinogram. This complete sinogram was then used to assemble the Fourier space and to interpolate the missing coefficients using DFR, see Figure 6.21.

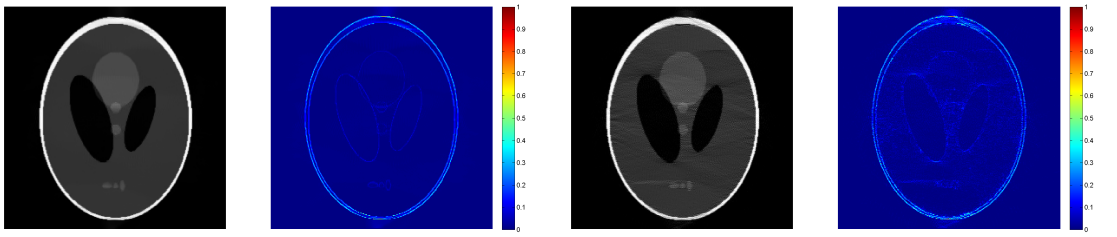


(a) *Original incomplete sinogram* ***(b)*** *Estimated missing projections* ***(c)*** *Gap-filled sinogram* ***(d)*** *DFR*

Figure 6.21: *DFR method from a gap-filled sinogram using OSEM reconstruction. The missing projections in (b) are obtained by re-projecting an OSEM estimate, and are used to fill the sinogram in (a). The estimated complete sinogram in (c) is used to obtain the reconstructed image (d) by DFR.*

The obtained image estimate is then used to obtain $\hat{y}_1$ for the CS reconstruction and the results are shown in Figure 6.22. Again, the best results were obtained

with $\hat{y}_2^{(0)} = 0$. Similarly to the previous methods, the best estimate was obtained in 7000 iterations. This approach gave the lowest %MSE of all tested reconstruction methods the and best results overall, see Figure 6.23. The image estimate shown in Figure 6.22 has almost completely corrected the missing wedge artifacts although the reconstruction seems to be slightly blurred by the filtering process. This is clear in the error images, where it can be seen that most of the error comes in the borders of the different regions. It can be seen that the image estimate (after 25000 iterations) presents some slight artifacts.



(a) 7000th iteration *(b) reconstruction error of (a)* *(c) 25000th iteration* *(d) reconstruction error of (c)*

Figure 6.22: *Results of applying CS algorithm directly over a DFR reconstruction from OSEM gap-filled sinogram.*

Regarding to the convergence in the %MSE values, inspecting Figure 6.23, after achieving a minimum %MSE value in the 7000th iteration, the curve converges to a higher value, but still better than all previously tested reconstructions.

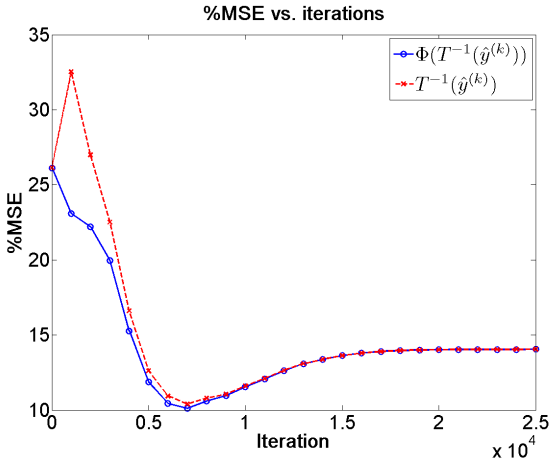
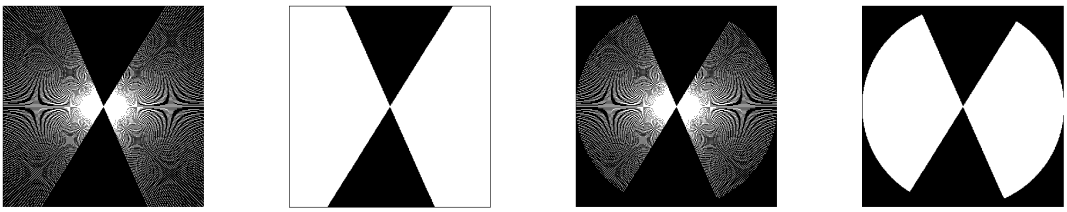


Figure 6.23: *%MSE results for CS algorithm using an inpainted DFR reconstruction.*

6.5.4 Variations of the Binary Mask

Four different versions of the sampling operator S, or binary mask, for the CS method were proposed and tested with the reconstruction methods to segment $\hat{y}_1$

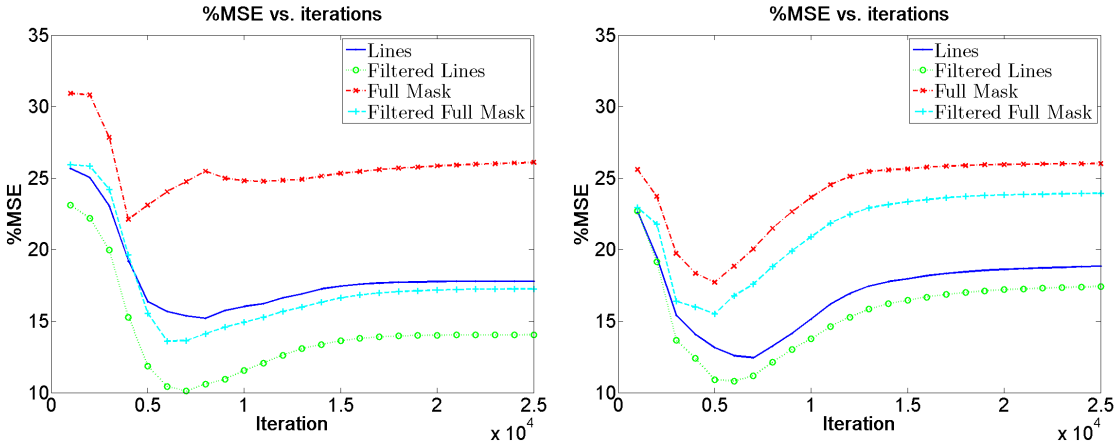
and $\hat{y}_2$. The masks used for the simulations are shown in Figure 6.24.



(a) lines *(b) full mask* *(c) filtered lines* *(d) filtered full mask*

Figure 6.24: *Variations of the Fourier space binary mask S for the angular range* $[-65°, +58°]$.

The %MSE curves obtained for the best two reconstruction methods are shown in Figure 6.25. The mask were used with the best two methods: CS with sinogram gap-filling using MLEM and CS with MLEM initialization. It can be seen in the curves that the best performance is always with the filtered version of the original mask, composed of filtered lines. The complete full mask performed worst in all cases.



(a) CS with sinogram gap-filling using MLEM *(b) CS with initialization by MLEM*

Figure 6.25: *Results for modifications of the binary mask S, %MSE for the best two reconstructions.*

Following the results shown in Figure 6.25, the selected version of S for the reconstructions was the filtered version of the original mask.

6.6 Real Data Reconstructions

The method proposed for the reconstruction of real Transmission Electron Microscope (TEM) data is the CS approach with initialization by MLEM. Due to the uniqueness and complexity of real biological data it is a difficult task to assess the reconstruction performance of the proposed method. Therefore, the evaluation of the reconstruction of real TEM data is performed visually, trying to examine the consistency of the interpolated data.

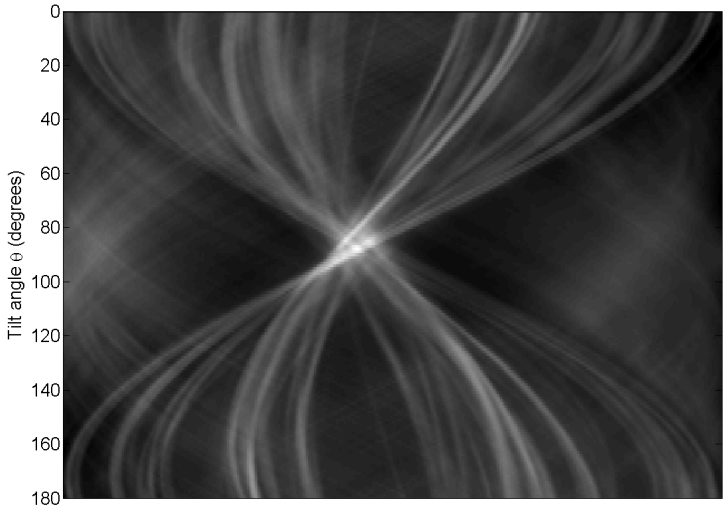


Figure 6.26: *Gap-filled 2D sinogram.*

Each of the 2D sinograms in the 3D stack of the acquired data were separately reconstructed to obtain the corresponding 2D slice of the reconstructed volume. Once all estimated slices are obtained, they are assembled together to form the 3D reconstructed volume. An example of a gap-filled sinogram with missing wedge correction is shown in Figure 6.26. It can be appreciated that the compensated missing wedge is visually coherent with the known part of the sinogram, and provides continuation to the sinusoidal trajectories described by the image features within the sinogram. It can be seen that for tilt angles close to 90°, the estimated electron counts are concentrated in the center of the radial axis of the sinogram. This is as expected since these projections correspond to tilts where the specimen slice is perpendicular to the electron beam.

A 3D representation of the filled sinogram stack is illustrated in Figure 6.27. It should be noted that the image features falling outside the field of view during tilting are not recovered during gap-filling, and this can be seen in the extreme regions of the sinogram. In Figure 6.28, a central slice of the reconstructed volumes using the proposed method and Weighted Back-projection are shown. They correspond

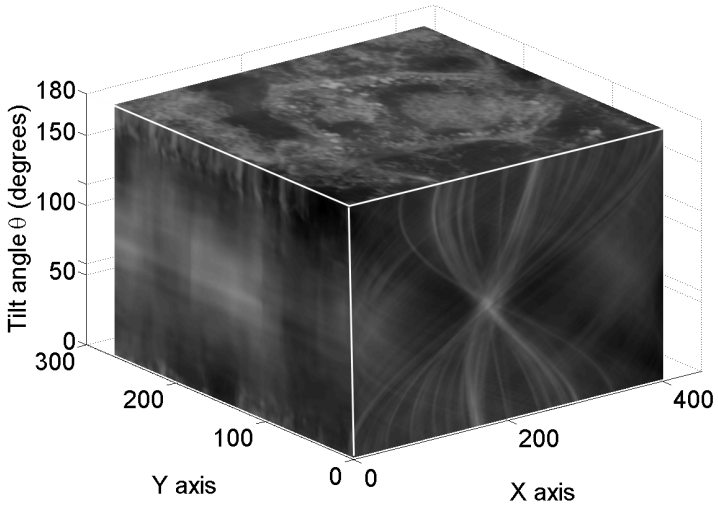
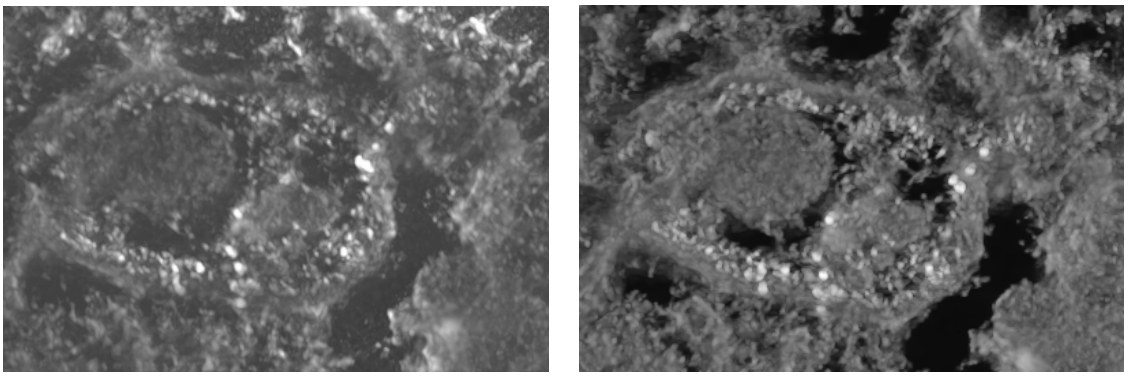


Figure 6.27: *Gap-filled 3D stack of sinograms.*

approximately to the same virtual cut in the volume although some slight differences in the rotation of the specimen. The gold particles in the proposed method appear to be defined better than in the WBP reconstruction.



(a) *Weighted Backprojection*

(b) *Reconstruction CS with MLEM initialization*

Figure 6.28: *Visual comparison of reconstruction methods.*

A different approach to evaluate visually the reconstruction is to perform a virtual tilting over the reconstructed specimen, simulating the tilting in the TEM acquisition. This can be seen in Figure 6.29. The specimen is virtually rotated and projections are acquired and shown. The projections corresponding to the missing wedge were reconstructed and the missing wedge does not exist anymore.

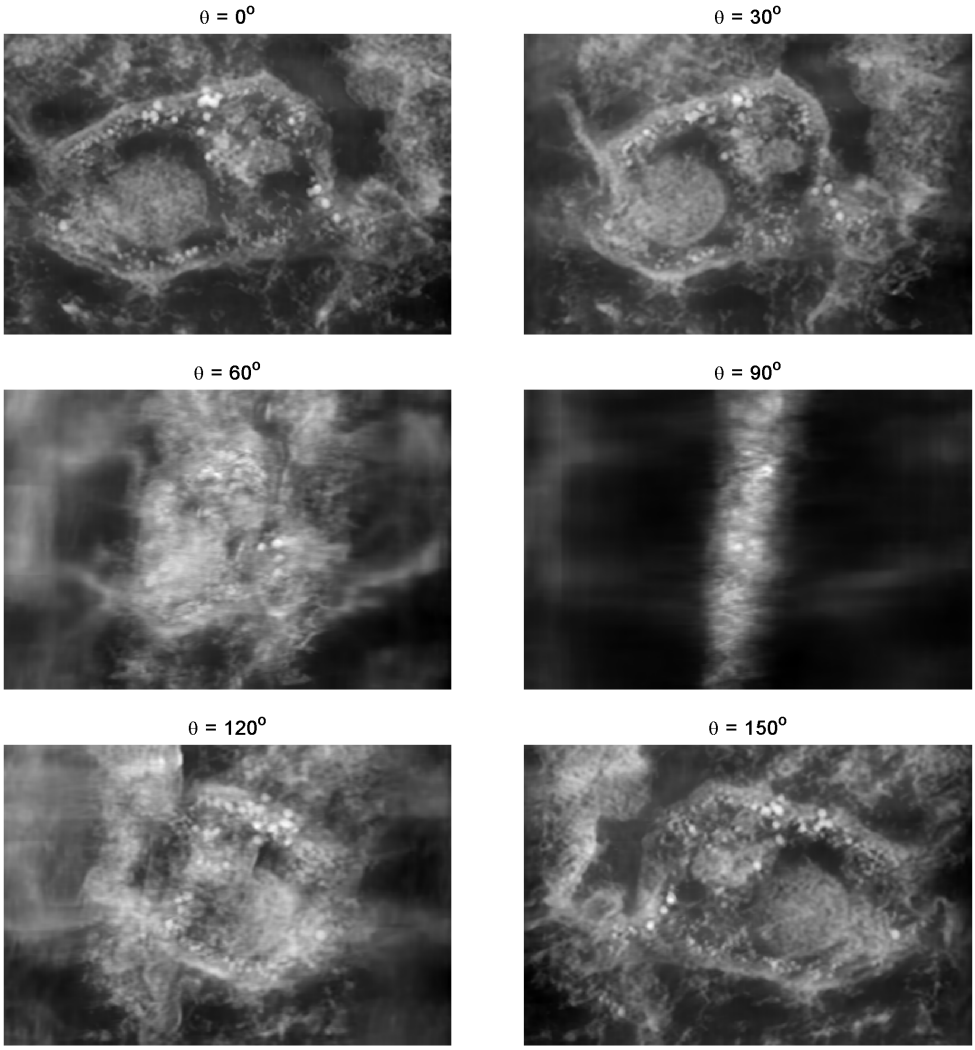


Figure 6.29: *Virtual tilting of the reconstructed specimen. Note that 30° and 90° tilt projections images are reconstructed projections within the missing wedge. The slab geometry of the slice can be seen in the 90° tilt.*

7. DISCUSSION AND CONCLUSION

This thesis focused on the limited-angle problem in Electron Tomography (ET), known as the missing wedge, which creates artifacts and anisotropic resolution in the reconstructed images. The main purpose was to obtain a reconstruction method capable of reducing the artifacts created by the missing wedge.

A common method to overcome the missing wedge problem is to use dual axis tilting or conical tilt in the Transmission Electron Microscope (TEM) acquisition stage [18, 24, 25, 28, 50]. This reduces the missing wedge to a missing pyramid and missing cone respectively, decreasing the amount of missing information and the reconstruction artifacts. However, this is not always possible and, therefore, other methods should be applied when dealing with single axis tilt ET.

A different approach to limited angle tomography consists of using *a priori* information to apply constraints in the reconstruction. This can be performed by using a Bayesian approach [54] to incorporate *a priori* probability densities, or by reconstruction based on projection onto convex sets (POCS) [56] where other constraints, such as shape and size, can be incorporated. However, when reconstructing unique complex biological specimens, wrong assumptions could lead to reconstruction artifacts, thus the use of powerful constraints is not possible. For this reason, a method that is not based on assumptions was needed, and thus other types of regularization would be preferred.

In this study, the 3D reconstruction was divided into a set of 2D reconstructions by organizing the TEM images into a stack, which consists of 2D sinograms. Each of these 2D sinograms was then used to reconstruct the corresponding slice of the volume.

The first approach for reconstruction with missing wedge correction addressed the problem directly in sinogram domain, trying to interpolate the missing gap. These methods were iterative projection back-projection, sinogram inpainting or gap-filling using a dedicated DCT filter. Iterative projection back-projection and gap-filling a dedicated DCT filter in Fourier space gave poor results and thus were rejected. Sinogram inpainting gave initially promising results, although the method failed with the presence of noise, and consequently was discarded.

The other methods were based on the Compressed Sensing (CS) approach, method that was proven to achieve perfect reconstruction in a simulated environment [64].

This approach consists of a stochastical reconstruction regularized in image domain by a spatially adaptive filter (BM3D). The BM3D filter does not need any assumptions of the data, which makes it a good choice for reconstructing unique complex specimens. However, the method presented in [64] is not directly applicable since it uses a set of known sparse Fourier coefficients from the original spectrum of the image, while our real TEM data is formed by projections of the specimen. That is why modifications to this method were necessary, focused on estimation of the Fourier coefficients from the available projections in the sinograms.

The modifications of the CS method approached the problem by using initialization by performing an initial reconstruction. An initial estimate in the form of preliminary image was obtained, and the Fourier coefficients needed for the CS approach could be segmented from the estimate's spectrum. In addition, modifications for the binary mask S, which segments the Fourier coefficients, were proposed. The proposed modifications for the CS method were:

- Initialization by Filtered Back-Projection (FBP)
- Initialization by Maximum Likelihood Expectation Maximization (MLEM)
- Initialization by interpolation in Fourier space and sinogram gap-filling using inpainting
- Initialization by interpolation in Fourier space and sinogram gap-filling using MLEM

All proposed methods improved the reconstruction compared to normal FBP reconstruction, although the best results were obtained by using MLEM initialization and by using interpolation in Fourier space of a MLEM gap-filled sinogram. The results obtained by CS approach with MLEM initialization were among the best obtained, and the missing wedge artifacts were considerably reduced in 7000 iterations. In terms of computational time, FBP initialization is better as the initialization requires less time than MLEM and the CS procedure achieves its best reconstruction earlier (in 1000 iterations). However, with FBP initialization the estimate is worse than the result with MLEM initialization.

The initialization by interpolation in Fourier space and sinogram gap-filling using inpainting gave good results although the method was discarded due to the poor performance of image-inpainting with real data. The best results overall were obtained with the CS method using initialization with interpolation in Fourier space of a MLEM gap-filled sinogram. The interpolation in Fourier space was performed by means of Direct Fourier Reconstruction (DFR). In this method, MLEM was used to obtain an estimate of the projections corresponding to the missing wedge, and these

projections were used to fill the original incomplete sinogram. The estimated complete sinogram was then used to estimate the complete Fourier space, using DFR, and the obtained image (and its spectrum) was then used to obtain an estimate of the portion of the spectrum used as the known part for the CS procedure. The unknown portion was initially set to zero. The obtained results were good, providing an accurate estimate of the original image, as shown in Figure 7.1.

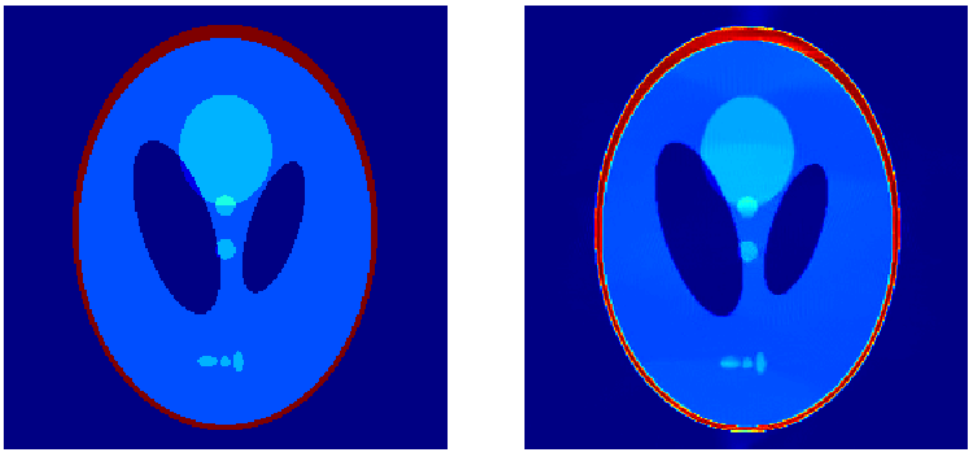


Figure 7.1: *Left: original head phantom. Right: Best reconstruction. Obtained with CS method using initialization with Direct Fourier Reconstruction of MLEM gap-filled sinogram.*

The chosen evaluation method, in addition to visual inspection, was the mean squared error (MSE) normalized by the MSE obtained in a reconstruction where no gap-filling was used. The images reconstructed with FBP were used as a reference. However, it must be noted that other assessment methods could be possible, where other measures, such as contrast, could be evaluated. For the proposed CS modifications, the %MSE curves showed similar convergence behavior. The best reconstructions were obtained within the first 7000 iterations, converging to poorer estimates in further iterations, with some artifacts visible in the reconstructed images. Further work should be done in order to assure convergence of the reconstruction to a better estimate, avoiding the appearance of new artifacts.

The reconstruction method was also applied to real Transmission Electron Microscope (TEM) data. The assessment of the reconstruction using real data was not straight-forward since the biological content of the original specimens were not known, there was no ground truth was available. Therefore, only visual evaluation was performed, evaluating the consistency of the interpolated data with the known available data. The estimated missing projections were consistent with the known projections in the sinogram, continuing the expected sinusoidal trajectories. Thus,

the 3D data was complete and no effect of the missing wedge could be detected.

An important factor in the reconstruction of TEM tomograms is alignment. Some of the problems faced during the reconstruction using real TEM data were because of the incorrect alignment of the projection images. If successive images are not perfectly aligned the reconstructed tomogram presents artifacts that could be mistaken as missing wedge artifacts, complicating the evaluation of the reconstruction. Furthermore, some of the methods used in this work, especially MLEM reconstruction, are quite sensitive to misalignments and could yield further reconstruction errors. Due to time limitations, this work did not go deep into alignment solutions although some initial attempts were made to correct possible misalignments.

Not only alignment between projections can cause complications, but also the center of rotation of the sinogram must be perfectly placed in the center of the sinogram. This is defined by the location of the tilting axis in the tilt series, which it is needed to be aligned with y axis in the center of the projection image. This condition is essential in order to be able to decompose the 3D reconstruction into a set of 2D reconstructions. Otherwise strong artifacts appear hindering the evaluation of the reconstructed volume.

In this study, the simulations with the numerical phantoms were performed by using noise-free projections, although the real TEM data was noisy. As future work, the reconstruction methods should be thoroughly studied in the presence of noise. TEM images present a low signal-to-noise ratio (SNR), so noise should be taken care of. A proposed solution is to use pre-filtering of the TEM projections. A promising possibility is the BM3D filter, which is one of the best filters available for filtering additive Gaussian noise. Therefore, provided that the predominant noise is additive Gaussian noise, the BM3D filter is the best choice for denoising the projections. In some cases, the presence of Poisson noise in the projection images should not be neglected, therefore, the performance of the BM3D filter should be also studied in these cases where a combination of both types of noise are present. A possible denoising example of a TEM image with the BM3D filter is shown in Figure 7.2

Finally, some words regarding computational complexity are needed. The simulations were performed with a Shepp-Logan head phantom of 256×256 pixels size, and took less than an hour with a normal PC. However, real TEM data is of considerably greater size. Assuming projection images are of size 1024×1024, which is a normal size, the 3D sinogram stack is composed of 1024 2D sinograms of size 1024×120 (assuming 120 tilt angles). In this manner, a reconstruction of a 1024×1024 pixels slice is repeated 1024 times, which takes days, even weeks. For this reason, downsized versions were reconstructed. Further work should include a computationally effective method in order to achieve affordable reconstruction times. Since the slices can be reconstructed separately, grid computing is easily applicable. In this man-

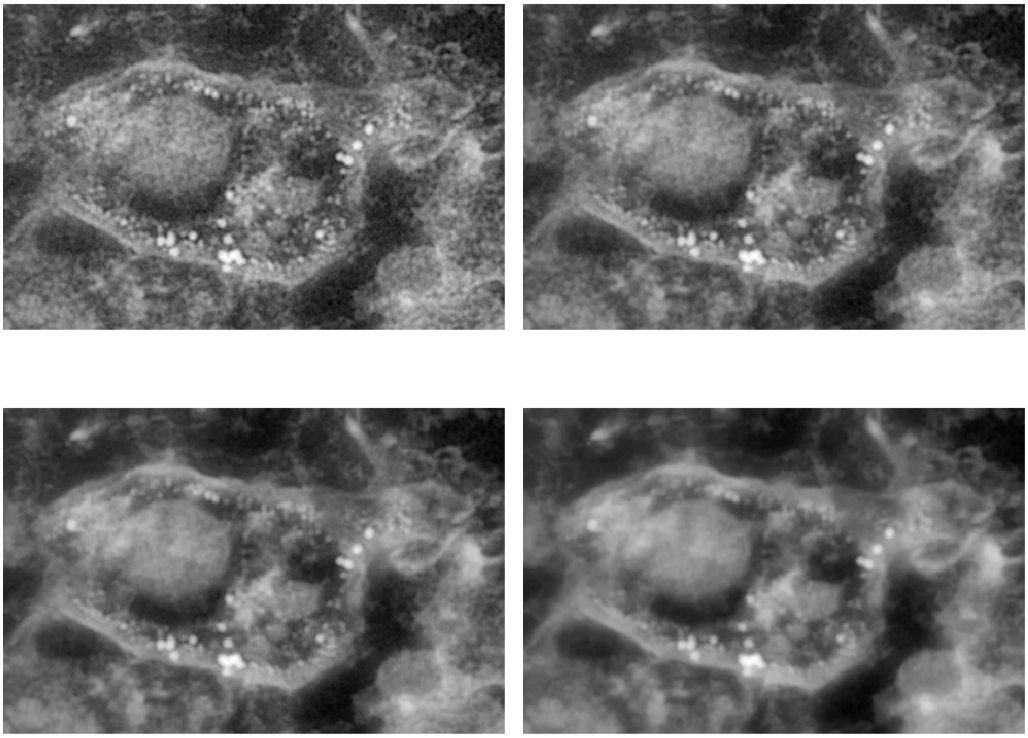


Figure 7.2: *Denoising of a noisy TEM projection image using BM3D for different denoising levels. The filter was applied iteratively, for 1 (top left), 10(top right), 20 (bottom left) and 30 iterations (bottom right).*

ner, subdividing the reconstruction into N subsets would decrease approximately N times the reconstruction time. Initial tests have been performed using the Techila Grid distributed computing environment [69], achieving successful and satisfactory results. By using Techila, and by distributing the reconstruction in separate slices, a complete 3D reconstruction could be performed in few hours, instead of few days needed in a single computer.

As a final conclusion, this study showed that is possible to correct the artifacts caused by the missing wedge in Electron Tomography. Reconstruction methods for limited angle tomography were proposed and applied to real TEM data, obtaining a reconstruction where the missing wedge is no longer existent.

REFERENCES

[1] J. Frank, "Introduction: Principles of electron tomography," in *Electron Tomography* (J. Frank, ed.), pp. 1–15, Springer New York, 2006.

[2] D. J. De Rosier and A. Klug, "Reconstruction of three dimensional structures from electron micrographs," *Nature*, vol. 217, no. 5124, pp. 130 – 134, 1968.

[3] P. A. Midgley and M. Weyland, "3d electron microscopy in the physical sciences: the development of z-contrast and eftem tomography," *Ultramicroscopy*, vol. 96, no. 3-4, pp. 413 – 431, 2003. Proceedings of the International Workshop on Strategies and Advances in Atomic Level Spectroscopy and Analysis.

[4] M. Marko, C.-E. Hsieh, and C. A. Mannella, "Electron tomography of frozen-hydrated sections of cells and tissues," in *Electron Tomography* (J. Frank, ed.), pp. 49–81, Springer New York, 2006.

[5] W. Baumeister, R. Grimm, and J. Walz, "Electron tomography of molecules and cells," *Trends in Cell Biology*, vol. 9, no. 2, pp. 81 – 85, 1999.

[6] R. Egerton, "An introduction to microscopy," in *Physical Principles of Electron Microscopy*, pp. 1–25, Springer US, 2005.

[7] J. M. Plitzko and W. Baumeister, "Cryoelectron tomography (cet)," in *Science of Microscopy* (P. W. Hawkes and J. C. H. Spence, eds.), pp. 535–604, Springer New York, 2007.

[8] W. Baumeister, "Electron tomography: towards visualizing the molecular organization of the cytoplasm," *Current Opinion in Structural Biology*, vol. 12, no. 5, pp. 679 – 684, 2002.

[9] D. B. Williams, C. B. Carter, D. B. Williams, and C. B. Carter, "The transmission electron microscope," in *Transmission Electron Microscopy*, pp. 3–22, Springer US, 2009.

[10] R. Egerton, *Physical Principles of Electron Microscopy. An Introduction to TEM, SEM, and AEM.* Springer US, 2005.

[11] D. B. Williams, C. B. Carter, D. B. Williams, and C. B. Carter, *Transmission Electron Microscopy.* Springer US, 2009.

[12] P. W. Hawkes and J. C. H. Spence, *Science of Microscopy.* Springer New York, 2007.

[13] R. Egerton, "The transmission electron microscope," in *Physical Principles of Electron Microscopy*, pp. 57–92, Springer US, 2005.

[14] JEOL, "Jeol jem-2100f transmission electron microscope." `http://www.jeol.com/PRODUCTS/ElectronOptics/TransmissionElectronMicroscopesTEM/200kV/JEM2100F/tabid/124/Default.aspx`, accessed in Feb. 2011.

[15] D. B. Williams, C. B. Carter, D. B. Williams, and C. B. Carter, "Scattering and diffraction," in *Transmission Electron Microscopy*, pp. 23–38, Springer US, 2009.

[16] P. W. Hawkes, "The electron microscope as a structure projector," in *Electron Tomography* (J. Frank, ed.), pp. 83–111, Springer New York, 2006.

[17] A. J. Koster and M. Bárcena, "Cryotomography: Low-dose automated tomography of frozen-hydrated specimens," in *Electron Tomography* (J. Frank, ed.), pp. 113–161, Springer New York, 2006.

[18] V. Lučić, F. Förster, and W. Baumeister, "Structural studies by electron tomography: From cells to molecules," *Annual Review of Biochemistry*, vol. 74, no. 1, pp. 833–865, 2005.

[19] M. F. Hohmann-Marriott, A. M. L. Meene, and R. W. Roberson, "Three-dimensional electron microscopy. a gateway to photosynthetic structure," in *Biophysical Techniques in Photosynthesis* (T. J. Aartsma and J. Matysik, eds.), vol. 26 of *Advances in Photosynthesis and Respiration*, pp. 13–34, Springer Netherlands, 2008.

[20] R. Hegerl and A. S. Frangakis, "Denoising of electron tomograms," in *Electron Tomography* (J. Frank, ed.), pp. 331–352, Springer New York, 2006.

[21] A. S. Frangakis, *Noise reduction and segmentation techniques developed for multidimensional electron microscopy of biological specimens*. PhD thesis, Max Planck Institute for Biochemistry Am Klopferspitz 18 D-82152 Martinsried Germany, 2001.

[22] Wikipedia, "Tem-single-tilt." `http://en.wikipedia.org/wiki/File:TEM-Single-tilt.svg`, accessed in Feb. 2011.

[23] P. A. Midgley, "Tomography using the transmission electron microscope," in *Handbook of Microscopy for Nanotechnology* (N. Yao and Z. L. Wang, eds.), pp. 601–627, Springer US, 2005.

[24] I. Arslan, J. R. Tong, and P. A. Midgley, "Reducing the missing wedge: High-resolution dual axis tomography of inorganic materials," *Ultramicroscopy*, vol. 106, no. 11-12, pp. 994 – 1000, 2006. Proceedings of the International Workshop on Enhanced Data Generated by Electrons.

[25] S. Lanzavecchia, F. Cantele, P. Bellon, L. Zampighi, M. Kreman, E. Wright, and G. Zampighi, "Conical tomography of freeze-fracture replicas: a method for the study of integral membrane proteins inserted in phospholipid bilayers," *Journal of Structural Biology*, vol. 149, no. 1, pp. 87 – 98, 2005.

[26] R. I. Koning and A. J. Koster, "Cryo-electron tomography in biology and medicine," *Annals of Anatomy - Anatomischer Anzeiger*, vol. 191, no. 5, pp. 427 – 445, 2009.

[27] N. Kawase, M. Kato, H. Nishioka, and H. Jinnai, "Transmission electron microtomography without the missing wedge for quantitative structural analysis," *Ultramicroscopy*, vol. 107, no. 1, pp. 8 – 15, 2007.

[28] P. A. Penczek and J. Frank, "Resolution in electron tomography," in *Electron Tomography* (J. Frank, ed.), pp. 307–330, Springer New York, 2006.

[29] M. Radermacher and W. Hoppe, "Properties of 3-d reconstruction from projections by conical tilting compared to single-axis tilting," in *7h European Congress on Electron Microscopy* (D. Haag, ed.), pp. 132–133, 1980.

[30] P. van der Heide, X.-P. Xu, B. J. Marsh, D. Hanein, and N. Volkmann, "Efficient automatic noise reduction of electron tomographic reconstructions based on iterative median filtering," *Journal of Structural Biology*, vol. 158, no. 2, pp. 196 – 204, 2007. Structural Analysis of Supramolecular Assemblies by Hybrid Methods.

[31] A. Kobayashi, T. Fujigaya, M. Itoh, T. Taguchi, and H. Takano, "Technical note: A tool for determining rotational tilt axis with or without fiducial markers," *Ultramicroscopy*, vol. 110, no. 1, pp. 1 – 6, 2009.

[32] J. Jan, *Medical Image Processing, Reconstruction and Restoration: Concepts and Methods*. Boca Raton, Florida: CRC press, 2006.

[33] B. F. Hayden, "Slice recontruction." `http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0405/HAYDEN/Slice_Reconstruction.html`, accessed in Feb. 2011.

[34] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, N.J.: Prentice Hall, 2010.

[35] J.-M. Carazo, G. T. Herman, C. O. S. Sorzano, and R. Marabini, "Algorithms for three-dimensional reconstruction from the imperfect projection data provided by electron microscopy," in *Electron Tomography* (J. Frank, ed.), pp. 217–243, Springer New York, 2006.

[36] D. Zosso, M. Bach Cuadra, and J.-P. Thiran, "Direct fourier tomographic reconstruction image-to-image filter," 2007.

[37] H. Stark, J. Woods, I. Paul, and R. Hingorani, "Direct fourier reconstruction in computer tomography," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, pp. 237 – 245, Apr. 1981.

[38] J. Walden, "Analysis of the direct fourier method for computer tomography," *Medical Imaging, IEEE Transactions on*, vol. 19, no. 3, pp. 211 –222, 2000.

[39] D. Gottleib, B. Gustafsson, and P. Forssen, "On the direct fourier method for computer tomography," *Medical Imaging, IEEE Transactions on*, vol. 19, no. 3, pp. 223 –232, 2000.

[40] B. M. W. Tsui and E. C. Frey, "Analytic image reconstruction methods in emission computed tomography," in *Quantitative Analysis in Nuclear Medicine Imaging* (H. Zaidi, ed.), pp. 82–106, Springer US, 2006.

[41] M. Radermacher, "Weighted back-projection methods," in *Electron Tomography* (J. Frank, ed.), pp. 245–273, Springer New York, 2006.

[42] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society of Industrial and Applied Mathematics, 2001.

[43] J.-J. Fernández, A. F. Lawrence, J. Roca, I. García, M. H. Ellisman, and J.-M. Carazo, "High-performance electron tomography of complex biological specimens," *Journal of Structural Biology*, vol. 138, no. 1-2, pp. 6 – 20, 2002.

[44] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, vol. 39, no. 1, pp. 1–38, 1977.

[45] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *Medical Imaging, IEEE Transactions on*, vol. 1, no. 2, pp. 113 –122, 1982.

[46] J. Ollinger, "Maximum-likelihood reconstruction of transmission images in emission computed tomography via the em algorithm," *Medical Imaging, IEEE Transactions on*, vol. 13, pp. 89 –101, Mar. 1994.

[47] S. Alenius, "Reconstruction from projections." `http://www.cs.tut.fi/sgn/m2obsi/m2obsiWWW/tomoRGI.pdf`, accessed in Apr. 2011.

[48] H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *Medical Imaging, IEEE Transactions on*, vol. 13, pp. 601 –609, Dec. 1994.

[49] S. Alenius and U. Ruotsalainen, "Generalization of median root prior reconstruction," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 11, pp. 1413 –1420, 2002.

[50] S. Lanzavecchia and P. L. Bellon, "Electron tomography in conical tilt geometry. the accuracy of a direct fourier method (dfm) and the suppression of non-tomographic noise," *Ultramicroscopy*, vol. 63, no. 3-4, pp. 247 – 261, 1996.

[51] K. C. Tam and V. Perez-Mendez, "Tomographical imaging with limited-angle input," *J. Opt. Soc. Am.*, vol. 71, pp. 582–592, May 1981.

[52] J. Carazo *et al.*, "Discrete tomography in electron microscopy," in *Discrete Tomography: Foundations, Algorithms, and Applications* (G. T. Herman and A. Kuba, eds.), pp. 405–416, Birkhäuser, 1999.

[53] S. F. Yau and S. H. Wong, "A linear sinogram extrapolator for limited angle tomography," in *Signal Processing, 1996., 3rd International Conference on*, vol. 1, pp. 386 –389 vol.1, Oct. 1996.

[54] K. M. Hanson and G. W. Wecksung, "Bayesian approach to limited-angle reconstruction in computed tomography," *J. Opt. Soc. Am.*, vol. 73, pp. 1501–1509, Nov 1983.

[55] I. Kazantsev, M. Samuel, and R. Lewitt, "Limited angle tomography and ridge functions," in *Nuclear Science Symposium Conference Record, 2002 IEEE*, vol. 3, pp. 1706 – 1710 vol.3, 2002.

[56] H. Kudo and T. Saito, "Sinogram recovery with the method of convex projections for limited-data reconstruction in computed tomography," *J. Opt. Soc. Am. A*, vol. 8, pp. 1148–1160, Jul 1991.

[57] S. S. Brandt, "Markerless alignment in electron tomography," in *Electron Tomography* (J. Frank, ed.), pp. 187–215, Springer New York, 2006.

[58] K. C. Tam and V. Perez-Mendez, "Limited-angle 3-d reconstructions usin fourier transform iterations and radon transform iterations," *International Optical Computing Conference*, April 1980.

[59] U. Tuna, S. Peltonen, and U. Ruotsalainen, "Data estimation for the ecat hrrt sinograms by utilizing the dct domain," in *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pp. 5076 –5080, 2008.

[60] U. Tuna, S. Peltonen, and U. Ruotsalainen, "Gap-filling for the high-resolution pet sinograms with a dedicated dct-domain filter," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 3, pp. 830 –839, 2010.

[61] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, (New York, NY, USA), pp. 417–424, ACM Press/Addison-Wesley Publishing Co., 2000.

[62] B. S. J. S. Loukiala A, Tuna U and R. U., "Gap-filling methods for the 3d plantis data," *Physics in Medicine and Biology*, vol. 55, no. 20, pp. 6125–6139, 2010.

[63] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289 –1306, 2006.

[64] K. Egiazarian, A. Foi, and V. Katkovnik, "Compressed sensing image reconstruction via recursive spatially adaptive filtering," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 1, pp. I –549 –I –552, 2007.

[65] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3d filtering," in *Electronic Imaging 2006, Proc. SPIE 6064,*, no. 6064A-30, 2006.

[66] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080 –2095, 2007.

[67] L. A. Shepp and B. F. Logan, "Reconstructing interior head tissue from x-ray transmissions," *Nuclear Science, IEEE Transactions on*, vol. 21, no. 1, pp. 228 –236, 1974.

[68] K. Egiazarian, A. Foi, V. Katkovnik, and A. Danielyan, "Compressed sensing image reconstruction via recursive spatially adaptive filtering, nonparametric stochastic approximation approach." `http://www.cs.tut.fi/~comsens/`, accessed in Mar. 2011.

[69] Techila Technologies Ltd, "Techila grid." `http://www.techila.fi/`, accessed in Jun. 2011.

[70] Y. Liu, P. A. Penczek, B. F. McEwen, and J. Frank, "A marker-free alignment method for electron tomography," *Ultramicroscopy*, vol. 58, no. 3-4, pp. 393 – 402, 1995.

[71] S. Brandt, J. Heikkonen, and P. Engelhardt, "Automatic alignment of transmission electron microscope tilt series without fiducial markers," *Journal of Structural Biology*, vol. 136, no. 3, pp. 201 – 213, 2001.

[72] H. Winkler and K. A. Taylor, "Accurate marker-free alignment with simultaneous geometry determination and reconstruction of tilt series in electron tomography," *Ultramicroscopy*, vol. 106, no. 3, pp. 240 – 254, 2006.

[73] D. Castaño-Díez, A. Al-Amoudi, A.-M. Glynn, A. Seybert, and A. S. Frangakis, "Reprint of "fiducial-less alignment of cryo-sections" [j. struct. biol. 159 (2007) 413-423]," *Journal of Structural Biology*, vol. 161, no. 3, pp. 249 – 259, 2008. The 4th International Conference on Electron Tomography.

8. APPENDIX - DIRECT FOURIER RECONSTRUCTION ALGORITHM

The method of Direct Fourier Reconstruction (DFR) used is based on the algorithm described in [36], shown in Algorithm 1. The original method uses b-splines of order n_b, denoted by β^{n_b}, to interpolate radial samples in Fourier space and angular linear interpolation. However, the implemented version for this work uses bilinear interpolation. The parameters needed in Algorithm 1 are:

- α, the angular range covered
- n_z, zero-padding factor
- n_g, 2D Discrete Fourier Transform (DFT) oversampling
- $f_c \in [0, 1]$, normalized radial cut-off frequency
- x_0, y_0, m and n, output image offset and size

The chosen values for the reconstructions performed are $n_z = 2$, $n_g = 2$, $f_c = 1$, and $\alpha = 180$. An example of a reconstruction of a Shepp-Logan head phantom of size $m = 256, n = 256$ from 180 projections, is shown in Figure 8.1.

Figure 8.1: *Left: original head phantom. Right: DFR reconstruction from 180 projections. Images use the same colormap.*

Algorithm 1 Direct Fourier Reconstruction as presented in [36]

Require: Sinogram $p(r, \theta, z)$ with size $[c, b, a]$.
Ensure: $f(x, y, z)$

$b' \Leftarrow \lfloor 180 \cdot b/\alpha \rfloor$
$\Delta \Leftarrow 1 + 180 \cdot b/\alpha - b'$
$c' \Leftarrow c \cdot n_z, \quad c'' \Leftarrow c' \cdot n_g$
$r_{max} \Leftarrow f_c \cdot c'/2 - 1$
for $z = 0$ **to** $a - 1$ **do**
  for $\theta = 0$ **to** $b' - 1$ **do**
    for $r = 0$ **to** $c - 1$ **do**
      $r' \Leftarrow (r + c' - c/2) \operatorname{mod} c'$
      if $r' \operatorname{mod} 2 = 0$ **then**
        $p'(r') \Leftarrow p(r, \theta, z)$
      else
        $p'(r') \Leftarrow -p(r, \theta, z)$
      end if
    end for
    $P_\theta \Leftarrow FFT(p')$
  end for
  for $u, v = 0$ **to** $c'' - 1$ **do**
    $u' \Leftarrow u - c''/2, \quad v' \Leftarrow v - c''/2$
    $r'' \Leftarrow \sqrt{u'^2 + v'^2}/n_g$
    if $r'' < r_{max}$ **then**
      $\Theta \Leftarrow \tan_2^{-1}(v', u')$
      if $\Theta < 0$ **then**
        $\Theta \Leftarrow \Theta + \pi$
        $r'' \Leftarrow -r''$
      end if
      $\theta \Leftarrow b \cdot \Theta/\pi$
      if $\lfloor \theta \rfloor + 1 < b'$ **then**
        $F(u, v) \Leftarrow (1 - (\theta - \lfloor \theta \rfloor))\, \beta^{n_b}\left[P_{\lfloor \theta \rfloor}(r'' + c'/2)\right] + (\theta - \lfloor \theta \rfloor)\, \beta^{n_b}\left[P_{\lfloor \theta \rfloor + 1}(r'' + c'/2)\right]$
      else
        $F(u, v) \Leftarrow (1 - \Delta^{-1}(\theta - \lfloor \theta \rfloor))\, \beta^{n_b}\left[P_{\lfloor \theta \rfloor}(r'' + c'/2)\right] + \Delta^{-1}(\theta - \lfloor \theta \rfloor)\, \beta^{n_b}\left[P_{\lfloor \theta \rfloor + 1 - b}(c'/2 - r'')\right]$
      end if
    end if
  end for
  $f' \Leftarrow FFT^{-1}(F)$
  for $x = x_0$ **to** $x_0 + m - 1$, $y = y_0$ **to** $y_0 + n - 1$ **do**
    $x' = x + c(n_z \cdot n_g - 1/2) \operatorname{mod} c'', \quad y' = y + c(n_z \cdot n_g - 1/2) \operatorname{mod} c'',$
    if $(x' + y') \operatorname{mod} 2 = 0$ **then**
      $f(x, y, z) = f'(x', y')$
    else
      $f(x, y, z) = -f'(x', y')$
    end if
  end for
end for

A. APPENDIX - ALIGNMENT OF PROJECTIONS

Misalignment is caused by the mechanical imperfections of the acquisition system in the Transmission Electron Microscope (TEM). The errors present are translational misalignment, rotational misalignment, image warping, skewness, rotation and magnification changes [26]. During acquisition most of these errors are rectified but some residual errors are still present, are must be corrected. That is why, before reconstructing the tomogram, some pre-processing is needed.

The imperfections present in the TEM images can be expressed mathematically making the *affine camera model* assumption [57]. This model expresses the projections as an affine transformation from a 3D coordinate system into a 2D one. Thus, the 3D vector $\mathbf{x}$ is projected onto the image plane, giving the 2D vector $\hat{\mathbf{m}}$:

$$\hat{\mathbf{m}} = \mathbf{A}\mathbf{x} + \mathbf{t}, \tag{A.1}$$

where the matrix $\mathbf{A}$ includes the rotations, projection operator and intrinsic parameters of the affine camera. The vector $\mathbf{t}$ represents the translation in the 2D image plane. In the case of single-tilt axis acquisition and neglecting minor effects this equation can be decomposed and further simplified

$$\hat{\mathbf{m}} = \mathbf{K}\mathbf{R}_\alpha\mathbf{x} + \mathbf{t} = s\mathbf{I_{2x2}}\mathbf{R}_\alpha\mathbf{x} + \mathbf{t}, \tag{A.2}$$

where $\mathbf{K}$ represent the intrinsic parameters of the affine camera and $\mathbf{R}_\alpha$ corresponds to the in-plane rotation of the tilt axis. The intrinsic parameters of the affine camera can be neglected, remaining only the scale factor s and then $\mathbf{K} = s\mathbf{I}_{2x2}$ as expressed in Equation A.2. Thus, in order to align the projections, these transformations need to be inverted, resulting in a 2D transformation where the aligned projections are [57]

$$\mathbf{m}_{\text{aligned}} = \mathbf{R}_\alpha^{\mathbf{T}}\mathbf{K}^{-\mathbf{1}}\left(\mathbf{m} - \mathbf{t}\right). \tag{A.3}$$

Thus, in order to have an aligned set of projections, neglecting possible affine transformations $\mathbf{K} = \mathbf{I_{2\times 2}}$, only the rotation $\mathbf{R}_\alpha$ and translational misalignment $\mathbf{t}$ need to be calculated. There are two main classical corrections for these two errors. For the former, Liu *et al.* proposed a method that solves rotational misalignments of

projections by making use of the so called common line [70]. This method makes use of the Fourier slice theorem, which states that all ideal projections at different tilt angles of an object share a common line in the 3D Fourier space [19]. Common line alignment uses this fact to find the line in Fourier space common to all projections in the data set. However, their method has been proven to be unreliable [57], and only in some determined cases works properly.

Usually, if the rotational deviations between views $\mathbf{R}_\alpha$ are negligible, only the translation errors $\mathbf{t}$ are needed to be corrected. For this case, an alignment method based on cross-correlation between slices is proven to be effective [57].

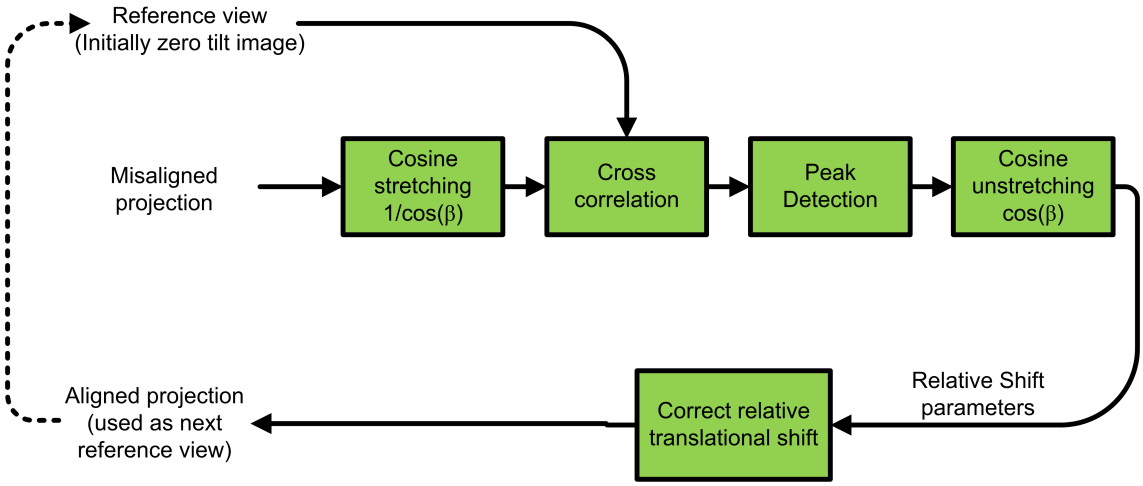
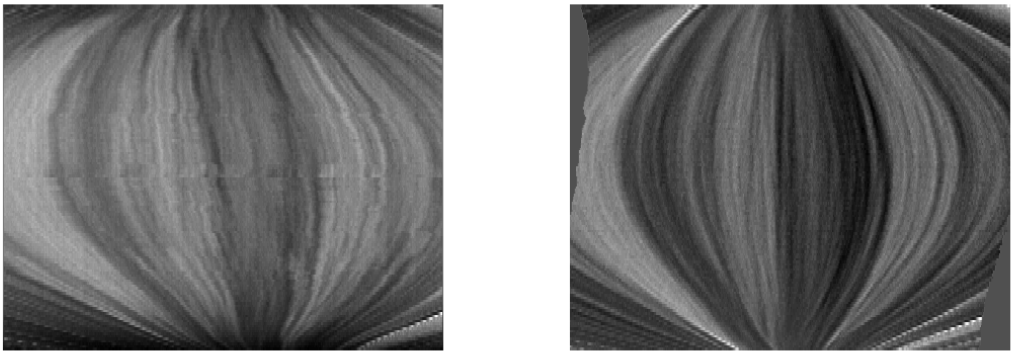


Figure A.1: *Recursive alignment procedure described in [57].*

The cross-correlation alignment, see Figure A.1, uses cross-correlation between successive pairs of projections to align the tilt series. Initially, the zero-tilt view is set as the reference view, and used to correct the neighboring tilt projection images. Before applying actual cross-correlation, it is recommended to apply stretching to the image to compensate the foreshortening due to specimen tilting [57]. The stretching is performed by a factor of $1/\cos\beta$, where β is the difference in the tilt angles between the two images. This is performed in the direction perpendicular to the tilt axis. Once the correlation is performed by usual methods (*e.g.* by means of Fast Fourier Transform), it is possible to estimate the offset between both projections by using peak detection. After performing cosine unstretching by a factor of $\cos\beta$ the translational parameters $\mathbf{t}$ are obtained and the projection can be aligned by using the transformation defined in equation A.3.

It is not possible to use the zero-tilt view as a reference for pairwise alignment of every projection since the difference would be too great to obtain a good cross-correlation. For this reason, images are aligned pairwise in a recursive manner. Each projection in the tilt series is aligned using the previously aligned projection, using the zero-tilt projection as the initial reference. An example of a 2D sinogram

from a 3D sinogram stack before and after alignment by this procedure is shown in Figure A.2.



(a) Not aligned *(b) Aligned*

Figure A.2: *Sinogram before and after alignments. Note that after alignment the sinusoidal trajectories are now well defined in the sinogram. Due to the displacements caused by alignment, some of the extreme radial samples at different angles have gone out of the sinogram.*

This method, however, poses a problem. Possible alignment errors are accumulated during the recursive pairwise alignments, leading to bigger alignment errors in high tilt images [57]. If other errors, such as rotational changes, magnification changes and other imperfections are not negligible, more advanced methods should be used, see for instance [57, 71–73].